

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 21:27:56 ; Search time 7.74147 Seconds
(without alignments)
1007.182 Million cell updates/sec

Title: US-09-813-453A-4
Perfect score: 1343
Sequence: 1 MLITIDVGNHTVVLGLFDGE.....EPWLTMLGLRLVYERNVSRM 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	6.9	4928	4	US-09-036-987A-5
2	93	6.9	4928	4	US-09-370-700-5
3	91.5	6.8	1113	4	US-09-629-616-3
4	91	6.8	521	4	US-09-134-001C-5639
5	89	6.6	361	3	US-09-028-934-30
6	88.5	6.6	3724	2	US-08-804-227C-10
7	88.5	6.6	3724	2	US-08-804-198-4
8	87	6.5	234	4	US-09-134-001C-3362
9	86.5	6.4	361	2	US-08-729-214-25
10	86.5	6.4	361	3	US-09-028-934-25
11	85	6.3	731	2	US-08-911-364-1
12	85	6.3	733	4	US-08-464-700-2
13	85	6.3	3567	2	US-07-642-734C-4
14	85	6.3	3567	3	US-08-439-009A-4
15	84	6.3	549	3	US-08-886-886-13
16	83.5	6.2	3519	4	US-09-428-517-2
17	83.5	6.2	4150	4	US-09-428-517-2
18	83	6.2	201	2	US-08-911-364-2
19	83	6.2	330	4	US-09-134-001C-3582
20	81.5	6.1	432	2	US-08-677-049-8
21	81.5	6.1	792	2	US-08-678-039A-40
22	80.5	6.0	819	4	US-09-651-656-15
23	80.5	6.0	819	4	US-09-650-855-15
24	80	6.0	355	4	US-08-818-112-79
25	80	6.0	355	4	US-08-818-111-80
26	80	6.0	355	4	US-09-056-556-79
27	80	6.0	355	4	US-09-072-596-80

28 80 6.0 769 3 US-09-320-878-12 Sequence 12, Appl
29 80 6.0 809 4 US-09-105-537-24 Sequence 24, Appl
30 80 6.0 866 3 US-09-040-843-2 Sequence 2, Appl
31 80 6.0 866 4 US-09-621-855-2 Sequence 2, Appl
32 80 6.0 3782 4 US-09-105-537-4 Sequence 4, Appl
33 80 6.0 4545 2 US-08-804-227C-14 Sequence 14, Appl
34 80 6.0 4550 2 US-08-804-227C-8 Sequence 8, Appl
35 80 6.0 4550 2 US-08-804-198-2 Sequence 2, Appl
36 80 6.0 15281 2 US-08-471-119A-2 Sequence 2, Appl
37 79.5 5.9 263 4 US-09-134-001C-4512 Sequence 4512, Ap
38 79.5 5.9 493 4 US-09-177-349-5 Sequence 5, Appl
39 79.5 5.9 635 4 US-08-931-608A-5 Sequence 5, Appl
40 79.5 5.9 1529 2 US-08-728-470-10 Sequence 10, Appl
41 79.5 5.9 1529 4 US-08-719-641-10 Sequence 10, Appl
42 79.5 5.9 1600 2 US-08-617-697-10 Sequence 10, Appl
43 79 5.9 3739 3 US-09-320-878-2 Sequence 2, Appl
44 79 5.9 3739 4 US-09-105-537-33 Sequence 33, Appl
45 79 5.9 11877 4 US-09-105-537-6 Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-036-987A-5
; Sequence 5, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:

APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive

TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:

ADDRESSER: Dow AgroSciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4928 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-5

Query Match 6.9%; Score 93; DB 4; Length 4928;
Best Local Similarity 24.7%; Pred. No. 3.1;
Matches 54; Conservative 30; Mismatches 77; Indels 58; Gaps 11;

Db 96 --IG-----RLPAIIALVYALLPILRNTY-----TGIIKEVDPSLIEAAKGIGM- 137
QY 115 RIINAAVEL-----YGG--PAIVDFGATTFDAVSARGEYIGGVIAIGIEISVEA 165
Db 138 KPLRLTKVELPIAMPEVIMAGIRTAWLIGTATLAALIGAGG--LGDILLIGIDRNSA 195
QY 166 LGVKA-----QLRKE--VARPESVIGKNTV-----EAMQSGIVYGF 201
Db 196 LILIGAPALALAIIFDFILRWERSYKYLISLIGTIVIVIIIAIPMAQKGDKITF 255
QY 202 AGQV---DGVNRMARELADDDVTVIATGGLAPMVLGESSVI-----DEHEPWLTL 252
Db 256 AGKLGSEPSIITNMKILIEEDTDTVEVKDG-----MGKTSFLFNALKSDEIDGYLEFT 310
QY 253 G 253
Db 311 G 311

RESULT 5

3-09-028-934-30
; Sequence 30, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Philip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyroclitritin Biosynthesis Genes and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6117670artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/028,934
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1506/CIP7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-028-934-30

Query Match 6.6%; Score 89; DB 3; Length 361;
Best Local Similarity 22.6%; Pred. No. 0.17;

Matches 60; Conservative 35; Mismatches 93; Indels 78; Gaps 13;
QY 4 TIDVGNTHVTLGLFDEGDIHWRISTDSRTTDELAVALLOGLMGHPILGLDGLDIDG 63
Db 124 TADAQRSYT--GLGDEAHLESVRISMSLEAAIALTVELYDV-----PLRSPAFEEGCV 177
QY 64 IA--ICATVPSSLHLEVRTRYGVDPVAVLVPFGVKGT-----VPIL----- 104
Db 178 LAHLQKMWESIVYAIRFISPVQFYDELRFYEP--IRVGRSYLGPAGVEMPLEVLEHVL 236
QY 105 ----TDHP-----KEYGADRIINAIV--ELYGPAIVD--FGTATTFDAVSARGYIG 152
Db 237 WGSQSDHPAYLEFKETVLPVLPFAIRAIYARFAGQALVDRLVGEA---QAARERGPVG 293
QY 153 GVIAPGIEISVEALGVKQALRKIEVARPSVIGKNTVEAMQSGIVYGFAGQVDGVVNR 212
Db 294 AGLA-ALERIFEIL-----LHFRAPHLKLAERTYAAGGTG----- 327
QY 213 ARELADDDDDTVTIATGGLAPMVLGE 238
Db 328 -----PTIGSGGYAPSMGLD 342

RESULT 6

US-08-804-227C-10
; Sequence 10, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCI(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3724 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-227C-10

Query Match 6.6%; Score 88.5; DB 2; Length 3724;
Best Local Similarity 24.1%; Pred. No. 6.3;
Matches 77; Conservative 38; Mismatches 99; Indels 105; Gaps 19;
QY 2 LLTIDVGNTHVTLGLFDEGDIHWRISTDSRTTDELAVALLOGLMGHPILGLDGLDIDG 44
Db 983 LLIVPTGDQYTALA-----DTLEQAVASFGTVRRVAFDPARTGRAELFGLLETEINGDT 1037
QY 45 -----GLMG-----HPLLGLDELGDIDGIAICATVPSV--LHLEVRTRYRYGDVP 89

Db 1038 AVTGVVSLGLCTGRDPHAPV-----PVAVTATLALVQALADLGSTAPLWTVTCG 1088
QY 90 AVLVEPGVKTGPILTDHPKEVGADRI-INAAVAVEL---YGGPAIVVDFTG---ATTFD 142
Db 1089 AVATAP-----DELPCTAGAQLWGLGRVAALPEVWGG---LIDLPAAPDARVLD 1136
QY 143 AVSARGEYIGGVIAPGIE--ISVEALGVKGAQ-LRKIEVARPSRVIGKNTVEAMQSGIV 198
Db 1137 R-----LAGVLAEPGGEDQIAVRMAGVFGRRVLRNPNADSRPPAWRARGTVL----- 1182
QY 199 YGFAGOVGVNRMARELADD-----PDDVTVIATGGLAPMVL-----GE 238
Db 1183 --IAGDLTVPGRLVRSLLDGDADRVVLAGDPAPAAAAAGLTGVSVPVRCVDTDRAL 1240
QY 239 SSVIDEHEPWLTMGLRLV 257
Db 1241 AALLDEHAPTVAHVAPPLV 1259

ESULT 7

US-08-804-198-4

; Sequence 4, Application US/08804198

; Patent No. 5945320

; GENERAL INFORMATION:

; APPLICANT: Burgett, Stanley G.

; APPLICANT: Kuhstoss, Stuart A.

; APPLICANT: Rao, Nagaraja R.

; APPLICANT: Richardson, Mark A.

; APPLICANT: Rostock, Paul R., Jr.

; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PAUL R. CANTRELL 1138

; STREET: LILLY CORPORATE CENTER

; CITY: INDIANAPOLIS

; STATE: IN

; COUNTRY: USA

; ZIP: 46285

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: Macintosh

; OPERATING SYSTEM: Macintosh 7.0

; SOFTWARE: Microsoft Word 5.1

; CURRENT APPLICATION DATA:

; FILING DATE: US/08/804,198

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: CANTRELL, PAUL R.

; REGISTRATION NUMBER: 36,470

; REFERENCE/DOCKET NUMBER: P9113

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 317-276-3885

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3724 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

US-08-804-198-4

Query Match 6.6%; Score 88.5; DB 2; Length 3724;
Best Local Similarity 24.1%; Pred. No. 6.3;
Matches 77; Conservative 38; Mismatches 99; Indels 105; Gaps 19;

QY 2 LLTIDVGNTHVTLGFDGEDIVH-----WRISTDSRRTA-DELAVALLQ----- 44
Db 983 LLIVPTGQYDALA-----DTLEQAVASFGGTVRRVAFDPARTGRAELFGLLETEINGDT 1037
QY 45 -----GLMGM-----HPLLGLDELGDGIDGIAICATVPSV--LHELREVTTRYGDPV 89
Db 1038 AVTGVVSLGLCTGRDPHAPV-----PVAVTATLALVQALADLGSTAPLWTVTCG 1088

QY 90 AVLVEPGVKTGPILTDHPKEVGADRI-INAAVAVEL---YGGPAIVVDFTG---ATTFD 142
Db 1089 AVATAP-----DELPCTAGAQLWGLGRVAALPEVWGG---LIDLPAAPDARVLD 1136
QY 143 AVSARGEYIGGVIAPGIE--ISVEALGVKGAQ-LRKIEVARPSRVIGKNTVEAMQSGIV 198
Db 1137 R-----LAGVLAEPGGEDQIAVRMAGVFGRRVLRNPNADSRPPAWRARGTVL----- 1182
QY 199 YGFAGOVGVNRMARELADD-----PDDVTVIATGGLAPMVL-----GE 238
Db 1183 --IAGDLTVPGRLVRSLLDGDADRVVLAGDPAPAAAAAGLTGVSVPVRCVDTDRAL 1240
QY 239 SSVIDEHEPWLTMGLRLV 257
Db 1241 AALLDEHAPTVAHVAPPLV 1259

RESULT 8

US-09-134-001C-3362

; Sequence 3362, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCI

; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 3362

; LENGTH: 234

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3362

Query Match 6.5%; Score 87; DB 4; Length 234;
Best Local Similarity 24.7%; Pred. No. 0.15;
Matches 48; Conservative 27; Mismatches 69; Indels 50; Gaps 9;

QY 54 GDELGDGIDGIAICATVPSVLHRELREVTTRYGDPVAVLVEPGVKTGPILTDHPKEVGA 113
Db 66 GDIINFLEPATICFAIP--LYRKREVLKKY-----LQIFGGIAVGTII-----A 109
QY 114 DRIINAAVAVELYGGPAIVVDFTGATTFDAVSARGEYIGG-----VIAPGIEISVEA 165
Db 110 LLLIYLVAITTFQGNQIIASMLPQATTAIALPVSIGGIGVKELTSLAVILNAWVIS--A 167
QY 166 LGVKAQLRKIE--VARPSVIGKNTVEAMQSGIVTGFAGVDGVNRMARELADDDV 223
Db 168 LGAKIVLKLKISNPATR-----GALGTSGHTLGA--AAKELGETEESM 210
QY 224 TVIATGGLAPMVLG 237
Db 211 -----GSTAVVIGV 219

RESULT 9

US-08-729-214-25

; Sequence 25, Application US/08729214

; Patent No. 5817502

; GENERAL INFORMATION:

; APPLICANT: Ligon, James M.

; APPLICANT: Hill, Dwight Steven

; APPLICANT: Ryals, John Andrew

; APPLICANT: Hammer, Phillip E.

; APPLICANT: van Pee, Karl-Heinz

; APPLICANT: Kirner, Sabine

; TITLE OF INVENTION: Genes for the synthesis of

;; TITLE OF INVENTION: antipathogenic substances
;; NUMBER OF SEQUENCES: 27
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ciba-Geigy Corporation
;; STREET: 520 White Plains Road
;; CITY: Tarrytown
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10591
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/729,214
;; FILING DATE: TBA
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meigs, J. Timothy
;; REGISTRATION NUMBER: 38,241
;; REFERENCE/DOCKET NUMBER: CGC 1506/CIP5
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8587
;; TELEFAX: 919-541-8689
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 361 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-08-729-214-25

Query Match 6.4%; Score 86.5; DB 2; Length 361;
Best Local Similarity 22.9%; Pred. No. 0.32;
Matches 62; Conservative 35; Mismatches 99; Indels 75; Gaps 14;

QY 2 LITIDVGN-----THTVLGLFDGEDIVHWRISTDSRTADELAVLGLMGHPLIGD 55
DB 114 LHVTVMNPAADAQSRYSYGLRDEAHLESVRISMAALEAAIAVTVLSGV-----PLRSP 169
QY 56 ELGDGIDGIA--ICATVPSVLHREVTTRYGDPVAVLVEPGVKTG-----VP 102
DB 170 AFAQCDELAAYLQKVESVYAYRFISLOVFNELRPFYEP-IRVGGSYLPGAVEMP 228
QY 103 IL-----TDHP-----KEVGADRIINAVAL--ELYGGPAI---VVDFTATTFD 142
DB 229 LFVLEHVLWGSDHPAYREFKETYLPYVLPAYRAVYARFAGEPALVDRLD-----EVQ 283
QY 143 AVSARGEYIGGVIA---PGIEISVEALGVKGAQLKIEVARPSVIGKNTVEAMQSGIY 199
DB 284 AAGARGEYVAGLAALDPVFEV-----LLRFRAPHLKLAERAYEAGOSGP 330
QY 200 GFAGQV-DGVNRM-----RELADDP 220
DB 331 GSGGYAPSAVDLLALTRAARFLRALDEP 361

RESULT 10
US-09-028-934-25
; Sequence 25, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Philip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyrolytic Biosynthesis Genes and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 37

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: No. 6117670artis Corporation
;; STREET: 3054 Cornwallis Road
;; CITY: Research Triangle Park
;; STATE: NC
;; COUNTRY: USA
;; ZIP: 27709
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/028,934
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/729,214
;; FILING DATE: 09-OCT-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/258,261
;; FILING DATE: 08-JUN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meigs, J. Timothy
;; REGISTRATION NUMBER: 38,241
;; REFERENCE/DOCKET NUMBER: CGC1506/CIP7
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8587
;; TELEFAX: 919-541-8689
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 361 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-09-028-934-25

Query Match 6.4%; Score 86.5; DB 3; Length 361;
Best Local Similarity 22.9%; Pred. No. 0.32;
Matches 62; Conservative 35; Mismatches 99; Indels 75; Gaps 14;

QY 2 LITIDVGN-----THTVLGLFDGEDIVHWRISTDSRTADELAVLGLMGHPLIGD 55
DB 114 LHVTVMNPAADAQSRYSYGLRDEAHLESVRISMAALEAAIAVTVLSGV-----PLRSP 169
QY 56 ELGDGIDGIA--ICATVPSVLHREVTTRYGDPVAVLVEPGVKTG-----VP 102
DB 170 AFAQCDELAAYLQKVESVYAYRFISLOVFNELRPFYEP-IRVGGSYLPGAVEMP 228
QY 103 IL-----TDHP-----KEVGADRIINAVAL--ELYGGPAI---VVDFTATTFD 142
DB 229 LFVLEHVLWGSDHPAYREFKETYLPYVLPAYRAVYARFAGEPALVDRLD-----EVQ 283
QY 143 AVSARGEYIGGVIA---PGIEISVEALGVKGAQLKIEVARPSVIGKNTVEAMQSGIY 199
DB 284 AAGARGEYVAGLAALDPVFEV-----LLRFRAPHLKLAERAYEAGOSGP 330
QY 200 GFAGQV-DGVNRM-----RELADDP 220
DB 331 GSGGYAPSAVDLLALTRAARFLRALDEP 361

RESULT 11
US-08-911-364-1
; Sequence 1, Application US/08911364
; Patent No. 5969106
; GENERAL INFORMATION:
; APPLICANT: ROTHSTEIN, Aser
; APPLICANT: KEELY, Fred W.
; APPLICANT: ROTHSTEIN, Steven J.
; TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN
; TITLE OF INVENTION: ELASTIN AND OTHER FIBROUS PROTEINS
; NUMBER OF SEQUENCES: 8

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: FOLEY & LARDNER
;; STREET: 3000 K Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20007-5109
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/911,364
;; FILING DATE: 07-AUG-1997
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/023,552
;; FILING DATE: 07-AUG-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bent, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 041082/0104
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 672-5300
;; TELEFAX: (202) 672-5399
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 731 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
US-08-911-364-1

Query Match 6.3%; Score 85; DB 2; Length 731;
Best Local Similarity 24.0%; Pred. No. 1.4;
Matches 47; Conservative 13; Mismatches 86; Indels 50; Gaps 7;

QY 60 GIDGIAICATVPSVLHELREVTTRYGDV-----PAVLVEPGV-----KTG 100
DB 265 GIGGIAGVGTPTAAAAAAAATAAAGLVPGPGPGVGVGAGVPGVPGAG 324
QY 101 VPILTDH-----PKEVGADRIINAVAAVELYGG-----PAIVVD 134
DB 325 IPVPGAGIPGAAPGVVSPPEAAAKAAKAAKYGARPGVGVGGIPTYGVGAGGFFGFGV 384
QY 135 FGTATTFDAVSARGEY--IGGVIAPGIEISVEALGVKGAQLRKIEVARPSVIGKNTVEA 192
DB 385 VGGIPGVAGVPSVGVPGVGGV--PGVGLSPEAQAAAAAATAAAGLVPGPGVGVGAGGFFGFGV 442

QY 193 MQSGIVYGFAGQVDGV 208
DB 443 AQFGLVPG-VGVAPGV 457

RESULT 12
US-08-464-700-2
; Sequence 2, Application US/08464700
; Patent No. 6232458
; GENERAL INFORMATION:
; APPLICANT: WEISS, ANTHONY S
; APPLICANT: MARTIN, STEPHEN L
; TITLE OF INVENTION: SYNTHETIC POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/464,700
;; FILING DATE: 7-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: AU PL6520
;; FILING DATE: 22-DEC-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: AU PL9661
;; FILING DATE: 28-JUN-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/AU93/00655
;; FILING DATE: 16-DEC-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bak, Mary E.
;; REGISTRATION NUMBER: 31,215
;; REFERENCE/DOCKET NUMBER: GHC3USA
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 215-540-9200
;; TELEFAX: 215-540-5818
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 733 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
US-08-464-700-2

Query Match 6.3%; Score 85; DB 4; Length 733;
Best Local Similarity 24.0%; Pred. No. 1.4;
Matches 47; Conservative 13; Mismatches 86; Indels 50; Gaps 7;

QY 60 GIDGIAICATVPSVLHELREVTTRYGDV-----PAVLVEPGV-----KTG 100
DB 267 GIGGIAGVGTPTAAAAAAAATAAAGLVPGPGPGVGVGAGVPGVPGAG 326
QY 101 VPILTDH-----PKEVGADRIINAVAAVELYGG-----PAIVVD 134
DB 327 IPVPGAGIPGAAPGVVSPPEAAAKAAKAAKYGARPGVGVGGIPTYGVGAGGFFGFGV 386
QY 135 FGTATTFDAVSARGEY--IGGVIAPGIEISVEALGVKGAQLRKIEVARPSVIGKNTVEA 192
DB 387 VGGIPGVAGVPSVGVPGVGGV--PGVGLSPEAQAAAAAATAAAGLVPGPGVGVGAGGFFGFGV 444

QY 193 MQSGIVYGFAGQVDGV 208
DB 445 AQFGLVPG-VGVAPGV 459

RESULT 13
US-07-642-734C-4
; Sequence 4, Application US/07642734C
; Patent No. 5824513
; GENERAL INFORMATION:
; APPLICANT: Katz, L
; APPLICANT: Donadio, S
; APPLICANT: Mcalpine, J B
; TITLE OF INVENTION: Recombinant DNA Method for Producing
; TITLE OF INVENTION: Erythromycin Analogs
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward H. Gorman
; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
; CITY: Abbott Park
; STATE: IL
; COUNTRY: US
; ZIP: 60064-3500
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/642,734C
; FILING DATE: 17-JAN-91
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckers, Andreas M
; REGISTRATION NUMBER: 32652
; REFERENCE/DOCKET NUMBER: 4952.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-9396
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3567 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; S-07-642-734C-4

Query Match 6.3%; Score 85; DB 2; Length 3567;
Best Local Similarity 26.6%; Pred. No. 14;
Matches 69; Conservative 30; Mismatches 102; Indels 58; Gaps 16;

QY 20 EDIVHWRISTDSRRRTADELAVLLQGLMGHPLLDGLDGI--DGIAICATVPVSV---- 73
Db 3190 EDLGASAEIAACDTADRDALSALLDGL--PRPLTGVVHAAGVLADGLVTSIDPAVEQVL 3247
QY 74 -----LHELREVTTR-----YGDVPAVLVEPGVKTGPVILTDHPKEVGADRIINAV 120
Db 3248 RAKVDAANLHETLTANTGLSFVLFSSAASVLAGPG--QGV-----YAAANESLNAL 3297
QY 121 AAVELYGG-PAIVVDFTG-ATTFDAVSARGEYIG--GVIAPGIEISVEALGVKGALRK- 175
Db 3298 AALRRTRGLPAKALGWLWAQASEMTSGLDRIARTGVAALPTE---RALAFDSALRRG 3354
QY 176 IEVARP-----RSVIGK-NTVEAMQSGIVYG----FAGQVD-----GVNRMARELADDPDV 223
Db 3355 GEVFPPLSNRSALRAEFVPEVLRGMVRAKLRAAGAAAGPNVVDRLAGRSED---- 3410
QY 224 TVIATGGLAPMVLGESSVI 242
Db 3411 ---QVAGLAELVRSHAAV 3426

RESULT 14
; S-08-439-009A-4
; Sequence 4, Application US/08439009A
; Patent No. 6004787
; GENERAL INFORMATION:
; APPLICANT: Donadio, S
; APPLICANT: Katz, L
; APPLICANT: Mcalpine, J B
; TITLE OF INVENTION: Method of Directing Biosynthesis of
; TITLE OF INVENTION: Specific Polyketides
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven F. Weinstock
; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
; CITY: Abbott Park
; STATE: IL
; COUNTRY: US
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/439,009A
; FILING DATE: 11-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Casuto, Dianne
; REGISTRATION NUMBER: 40,943
; REFERENCE/DOCKET NUMBER: 4952.US.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3567 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-439-009A-4

Query Match 6.3%; Score 85; DB 3; Length 3567;
Best Local Similarity 26.6%; Pred. No. 14;
Matches 69; Conservative 30; Mismatches 102; Indels 58; Gaps 16;

QY 20 EDIVHWRISTDSRRRTADELAVLLQGLMGHPLLDGLDGI--DGIAICATVPVSV---- 73
Db 3190 EDLGASAEIAACDTADRDALSALLDGL--PRPLTGVVHAAGVLADGLVTSIDPAVEQVL 3247
QY 74 -----LHELREVTTR-----YGDVPAVLVEPGVKTGPVILTDHPKEVGADRIINAV 120
Db 3248 RAKVDAANLHETLTANTGLSFVLFSSAASVLAGPG--QGV-----YAAANESLNAL 3297
QY 121 AAVELYGG-PAIVVDFTG-ATTFDAVSARGEYIG--GVIAPGIEISVEALGVKGALRK- 175
Db 3298 AALRRTRGLPAKALGWLWAQASEMTSGLDRIARTGVAALPTE---RALAFDSALRRG 3354
QY 176 IEVARP-----RSVIGK-NTVEAMQSGIVYG----FAGQVD-----GVNRMARELADDPDV 223
Db 3355 GEVFPPLSNRSALRAEFVPEVLRGMVRAKLRAAGAAAGPNVVDRLAGRSED---- 3410
QY 224 TVIATGGLAPMVLGESSVI 242
Db 3411 ---QVAGLAELVRSHAAV 3426

RESULT 15
; US-08-886-886-13
; Sequence 13, Application US/08886886
; Patent No. 6107068
; GENERAL INFORMATION:
; APPLICANT: Katz, Leonard
; APPLICANT: Delcardayre, Stephen B.
; APPLICANT: Davies, Julian E.
; TITLE OF INVENTION: COENZYME A DISULFIDE REDUCTASE,
; TITLE OF INVENTION: AND INHIBITORS THEREOF USEFUL AS ANTIMICROBIAL
; TITLE OF INVENTION: AGENTS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/886,886
; FILING DATE: 02-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 22:08:46 ; Search time 12.0821 Seconds
(without alignments)
2373.336 Million cell updates/sec

Title: US-09-813-453A-4
Perfect score: 1343
Sequence: 1 MLTTDVGNTHTVLGLFDGE.....EPWLTMLGLRLYVNRVSRM 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1343	100.0	265	US-09-813-453A-4	Sequence 4, Appl
2	688.5	52.0	272	US-09-813-453A-5	Sequence 5, Appl
3	694.5	51.7	272	US-09-712-363-276	Sequence 276, App
4	663	49.4	258	US-09-813-453A-49	Sequence 49, Appl
5	662	49.3	254	US-09-813-453A-47	Sequence 47, Appl
6	655	48.8	258	US-09-813-453A-2	Sequence 2, Appl
7	643.5	47.9	256	US-09-813-453A-55	Sequence 55, Appl
8	638	47.5	262	US-09-813-453A-45	Sequence 45, Appl
9	629	46.8	255	US-09-813-453A-7	Sequence 7, Appl
10	554.5	41.3	233	US-09-813-453A-17	Sequence 17, Appl
11	527	39.2	258	US-09-813-453A-6	Sequence 6, Appl
12	517	38.5	260	US-09-813-453A-51	Sequence 51, Appl
13	498.5	37.1	250	US-09-813-453A-3	Sequence 3, Appl
14	492.5	36.7	219	US-09-813-453A-57	Sequence 57, Appl
15	409.5	30.5	262	US-09-813-453A-8	Sequence 8, Appl
16	342.5	25.5	246	US-09-813-453A-9	Sequence 9, Appl
17	295	22.0	257	US-09-813-453A-53	Sequence 53, Appl
18	291.5	21.7	273	US-09-813-453A-10	Sequence 10, Appl
19	280	20.8	212	US-09-813-453A-59	Sequence 59, Appl

20	210	15.6	249	9	US-09-813-453A-61	Sequence 61, Appl
21	205.5	15.3	241	9	US-09-813-453A-41	Sequence 41, Appl
22	184	13.7	241	9	US-09-813-453A-63	Sequence 63, Appl
23	181.5	13.5	262	9	US-09-813-453A-11	Sequence 11, Appl
24	180.5	13.4	592	9	US-09-813-453A-22	Sequence 22, Appl
25	179	13.3	460	9	US-09-813-453A-39	Sequence 39, Appl
26	177	13.2	592	9	US-09-813-453A-43	Sequence 43, Appl
27	176	13.1	249	9	US-09-813-453A-70	Sequence 70, Appl
28	174	13.0	237	9	US-09-813-453A-13	Sequence 13, Appl
29	139	10.3	242	9	US-09-813-453A-65	Sequence 65, Appl
30	138.5	10.3	248	9	US-09-813-453A-20	Sequence 20, Appl
31	134	10.0	267	9	US-09-813-453A-15	Sequence 15, Appl
32	128.5	9.6	229	9	US-09-813-453A-12	Sequence 12, Appl
33	95.5	7.1	472	9	US-09-712-363-202	Sequence 202, App
34	92.5	6.9	525	10	US-09-815-242-11952	Sequence 11952, A
35	92	6.9	221	9	US-09-738-626-4878	Sequence 4878, Ap
36	91.5	6.8	1113	9	US-09-738-626-5279	Sequence 5279, Ap
37	91.5	6.8	1113	9	US-10-284-138-3	Sequence 3, Appli
38	91.5	6.8	1113	9	US-10-284-334-3	Sequence 3, Appli
39	91.5	6.8	1113	10	US-09-836-470B-3	Sequence 3, Appli
40	90	6.7	379	9	US-09-975-139-8	Sequence 8, Appli
41	90	6.7	937	9	US-10-126-927-66	Sequence 66, Appl
42	90	6.7	937	9	US-10-126-931A-66	Sequence 66, Appl
43	89.5	6.7	794	10	US-09-815-242-5697	Sequence 5697, Ap
44	89.5	6.7	802	10	US-09-815-242-12668	Sequence 12668, A
45	88.5	6.6	1446	9	US-10-166-087-44	Sequence 44, Appl

ALIGNMENTS

RESULT 1

US-09-813-453A-4
; Sequence 4, Application US/09813453A
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: ANTIBIOTICS
; FILE REFERENCE: OGZ-001
; CURRENT APPLICATION NUMBER: US/09/813,453A
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/227,860
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/667,569
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 4
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor
US-09-813-453A-4

Query Match	100.0%	Score 1343;	DB 9;	Length 265;
Best Local Similarity	100.0%;	Pred. No. 1.4e-109;		
Matches 265;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MLTTDVGNTHTVLGLFDGEDIVHWRISTDSRTADELAVLQGLMGMPHLLGDELGDG	60	
DB	1	MLTTDVGNTHTVLGLFDGEDIVHWRISTDSRTADELAVLQGLMGMPHLLGDELGDG	60	
QY	61	IDGIAICATVPVSLHRELVTTRYGDPVAVLPVPGVKTGPILTDPKPEVGADRIINAV	120	
DB	61	IDGIAICATVPVSLHRELVTTRYGDPVAVLPVPGVKTGPILTDPKPEVGADRIINAV	120	
QY	121	AAVELYGGPAIVVDFTATTFFDAVSARGEYIGVIAPIGIEISVEALGVKAQLRKIEVAR	180	
DB	121	AAVELYGGPAIVVDFTATTFFDAVSARGEYIGVIAPIGIEISVEALGVKAQLRKIEVAR	180	
QY	181	PRSVTKNTVEAMQSGIVYGFAGQVDGVVNRMARLADDDPDVTYIATGGLAPMVLGESS	240	

```

Query Match          49.3%; Score 662; DB 9; Length 254;
Best Local Similarity 49.8%; Pred. No. 3.4e-50;
Matches 132; Conservative 55; Mismatches 64; Indels 14; Gaps 4;

1  MLLTIDVGNTHVVLGFLDGEDIVEHWRISTDSRTADELAVILQGLMGHPLIGDELG-- 58
   | : | | | | | | | | | : | : | | | : | : | | | : | : | : |
1  MTLVLDVGNNTVVLGYQDETLLVHHWRLATSRQKTEDEVAMTVRSLF-----DHAGLQ 53
   | | | | | : | | : | : | : | : | : | : | : | : | : | : |

59 -DGIDGIALCATVPSVLHRELREVTTRYGDVPAVLVEPGVKTGPILLTDHPKEVGADRII 117
   | | | | | : | | : | : | : | : | : | : | : | : | : | : |

54 FQDIDGIVLSSVPPMFMFSLEQMKKFIHWTP--MIIGPGIKTGLNIDKYNPKPEVGADRI 112
   | | | | | : | | : | : | : | : | : | : | : | : | : | : |

118 NAVAARELYGGPAIVVDFGTATTFDAVSARGEYIGGVIAPGISVEALGVKAOLRKIE 177
   | | | | | | | | | | | | | : : : | | | | | | | | | : : : | |
113 NAVAARELYGYPAIVVDFGTATTYCLINEKKQVAGGVIAPIGMISVEALYHRASKLPRIE 172
   | | | | | : | | : | : | : | : | : | : | : | : | : | : |

178 VARPSRVYKNTVEAMQSGIVTGYFAGQGVGVNRMARELADDDPDDVTIATGGLAPMVLG 237
   | : | : | | | | | | | : | | | | | : | | | | | : | | | | |
173 IAKPRQVGTNTIDSMQSGIFGYGVYQGVVYKMAQAESEP---KVIAATGGLAKLIGT 229
   | : | : | | | | | | | : | | | | | : | | | | | : | | | | |

238 ESSVIDEHPEWLTLMGLRLVYERNV 262
   | | | : : | | | | : | : | : | : | : | : | : | : | : |
230 ESETIDVIDSFELTKGLQLIYKKNV 254
   | | | : : | | | | : | : | : | : | : | : | : | : | : |

```

RESULT 7
 US-09-813-453A-55
 ; Sequence 55, Application US/09813453A
 ; Patent No. US20020168681A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yocum, R. Rogers
 ; APPLICANT: Patterson, Thomas A.
 ; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
 ; TITLE OF INVENTION: ANTIBIOTICS
 ; FILE REFERENCE: OGZ-001
 ; CURRENT APPLICATION NUMBER: US/09/813,453A
 ; CURRENT FILING DATE: 2001-03-20
 ; PRIOR APPLICATION NUMBER: US 60/227,860
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: 09/667,569
 ; PRIOR FILING DATE: 2000-09-21
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 55
 ; LENGTH: 256
 ; TYPE: PRT
 ; ORGANISM: Clostridium difficile
 US-09-813-453A-55
 Query Match 47.9%; Score 643.5; DB 9; Length 256;

Best Local Similarity 49.8%; Pred. No. 1.4e-48;
Matches 130; Conservative 53; Mismatches 71; Indels 7; Gaps 4;
QY 1 MLLTIDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDELGDG 60
Db 1 MLLVFDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDELGDG 56
QY 61 IDGTAICATVPSVLHRELVTRRYGDPVAVLVEPGVKTGPIILTDHPKEVGADRIINAV 120
Db 57 IDDVIISSVWPNVHSLFENFCIKYCKKOP-LIVGPIKTNIRKYNPNKQVGADRIINAV 115
QY 121 AAELVGGPAIVDFGATTTDAYSARGEYIGGVVAPLVEGVTGPIILTDHPKEVGADRIINAV 180
Db 116 AGIEKYGAPSLVDFGATTTDAYSARGEYIGGVVAPLVEGVTGPIILTDHPKEVGADRIINAV 175
QY 181 PRSVIGKNTVEMOSGIVYGFAGVDGVVNRMARLADDDPDVTVIATGGLAPMVLGESS 240
Db 176 PGMTICKSTVSAMOSGIIYGVGLVDKIISIMKEL--NCDDVKVIATGGLAKLIASETK 233
QY 241 VIDEHEPWLTLGRLLYERN 261
Db 234 SIDYVDGFLTEGLRIIYERN 254

RESULT 8
US-09-813-453a-45
; Sequence 45, Application US/09813453a
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; FILE REFERENCE: OGZ-001
; CURRENT APPLICATION NUMBER: US/09/813,453a
; PRIOR FILING DATE: 2001-03-20
; PRIOR FILING DATE: 2000-08-24
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-813-453a-45

Query Match 47.5%; Score 638; DB 9; Length 262;
Best Local Similarity 47.9%; Pred. No. 4.4e-48;
Matches 126; Conservative 55; Mismatches 74; Indels 8; Gaps 3;
QY 1 MLLTIDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDELGDG 60
Db 1 MIFVLDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDELGDG 56
QY 61 IDGTAICATVPSVLHRELVTRRYGDPVAVLVEPGVKTGPIILTDHPKEVGADRIINAV 120
Db 57 VKGIIVSVWPPIMFALERCMEKFKIKPLV-VGPIKTNIRKYNPNKQVGADRIINAV 115
QY 121 AAELVGGPAIVDFGATTTDAYSARGEYIGGVVAPLVEGVTGPIILTDHPKEVGADRIINAV 180
Db 116 AGHLVGLSPLIVDFGATTTDAYSARGEYIGGVVAPLVEGVTGPIILTDHPKEVGADRIINAV 175
QY 181 PRSVIGKNTVEMOSGIVYGFAGVDGVVNRMARLADDDPDVTVIATGGLAPMVLGESS 240
Db 176 PSSVVGKNTVEMOSGIIYGVGLVDKIISIMKEL--NCDDVKVIATGGLAKLIASETK 233
QY 241 VIDEHEPWLTLGRLLYERN 263
Db 233 VIDVDPFLTLKGLIYERN 255

RESULT 9
US-09-813-453a-7
; Sequence 7, Application US/09813453a
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; FILE REFERENCE: OGZ-001
; CURRENT APPLICATION NUMBER: US/09/813,453a
; PRIOR FILING DATE: 2001-03-20
; PRIOR FILING DATE: 2000-08-24
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Geobacter sulfurreducens
US-09-813-453a-7

Query Match 46.8%; Score 629; DB 9; Length 255;
Best Local Similarity 51.1%; Pred. No. 2.6e-47;
Matches 134; Conservative 42; Mismatches 76; Indels 10; Gaps 5;
QY 1 MLLTIDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDELGDG 59
Db 1 MLLVFDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDELGDG 55
QY 60 GIDGTAICATVPSVLHRELVTRRYGDPVAVLVEPGVKTGPIILTDHPKEVGADRIINAV 119
Db 56 QIRAVIISVWPPIMFALERCMEKFKIKPLV-VGPIKTNIRKYNPNKQVGADRIINAV 114
QY 120 AAELVGGPAIVDFGATTTDAYSARGEYIGGVVAPLVEGVTGPIILTDHPKEVGADRIINAV 179
Db 115 VAGYEKRTSLIVDFGATTTDAYSARGEYIGGVVAPLVEGVTGPIILTDHPKEVGADRIINAV 174
QY 180 RPRSVIGKNTVEMOSGIVYGFAGVDGVVNRMARLADDDPDVTVIATGGLAPMVLGESS 239
Db 175 RPSAIARTVNSMAGIYGVGLVDKIISIMKEL--NCDDVKVIATGGLAKLIASETK 231
QY 240 SVIDEHEPWLTLGRLLYERN 261
Db 232 KTIIEAVEYLTLEGLRIIYERN 253

RESULT 10
US-09-813-453a-17
; Sequence 17, Application US/09813453a
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; FILE REFERENCE: OGZ-001
; CURRENT APPLICATION NUMBER: US/09/813,453a
; PRIOR FILING DATE: 2001-03-20
; PRIOR FILING DATE: 2000-08-24
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-813-453a-17

Query Match 41.3%; Score 554.5; DB 9; Length 233;
Best Local Similarity 50.5%; Pred. No. 7.4e-41;
Matches 111; Conservative 39; Mismatches 65; Indels 5; Gaps 2;
QY 1 MLLTIDVGNTHVTLGLFDGEDIVEHWRISTDSRRRTADELAVLLQGLMGHPLGLGDELGDG 60
Db 1 MLLVIDVGNTHVTLGLFDGEDIVEHWRISTDSRRRTADELAVLLQGLMGHPLGLGDELGDG 60
QY 61 IDGTAICATVPSVLHRELVTRRYGDPVAVLVEPCVGTGVPILTDHPKEVGADRLINAV 120
Db 57 IDGLIISVVPIMPALERMCTKYEHIPEQI-VGPGMKTGLNIKYDNPKEVGADRLINAV 115
QY 121 AAVELYGGPAIVDFGTATTDAVSARGEYIGGVAPLVEPCVGTGVPILTDHPKEVGADRLINAV 180
Db 116 AAHLHLYGNPLIVDFGTATTTCYIDENKQYMGGAIPAGTITSEALYSRAAKLPRIETIR 175
QY 181 PRSVIGKNTVAMQSGIVYGFAGQVDGVNRMARELADDP 220
Db 176 PDNIIGKNTVAMQSGILFYGVGVGIVKRMKWAQKQDP 215

Query Match 39.2%; Score 527; DB 9; Length 258;
Best Local Similarity 45.6%; Pred. No. 2.1e-38;
Matches 119; Conservative 41; Mismatches 93; Indels 8; Gaps 3;
QY 1 MLLTIDVGNTHVTLGLFDGEDIVEHWRISTDSRRRTADELAVLLQGLMGHPLGLGDELGDG 60
Db 1 MLLCIDCGNTNTVFSVMDGTDFATWRIATDHRRTADEYFVWLNTLMQLKLGQ-----R 55
QY 61 IDGTAICATVPSVLHRELVTRRYGDPVAVLVEPCVGTGVPILTDHPKEVGADRLINAV 120
Db 56 ISEALISSTAPRVFNLCNRVDFCRPYVVGKPCELPAPRVDPGTGTVGPDRLVNTV 115
QY 121 AAVELYGGPAIVDFGTATTDAVSARGEYIGGVAPLVEPCVGTGVPILTDHPKEVGADRLINAV 180
Db 116 AGYDRHGGDLIVDEGTATTEDVAPDGAIGGVAPLVEPCVGTGVPILTDHPKEVGADRLINAV 175
QY 181 PRSVIGKNTVAMQSGIVYGFAGQVDGVNRMARELADDPDDVTVIATGGLAPMYLGESS 240
Db 176 PQGVIGTNTVACIQSGVYGVYIGLVGIVQIRME-RDRP--MKVIATGGLASLFDLGF 232
QY 241 VIDEHEPWLTLMLGLRLVYERN 261
Db 233 LFDKVEDDLTMHGLRLIFDYN 253

Query Match 37.1%; Score 498.5; DB 9; Length 250;
Best Local Similarity 43.3%; Pred. No. 6.2e-36;
Matches 104; Conservative 50; Mismatches 73; Indels 13; Gaps 4;
US-09-813-453A-51
; Sequence 51, Application US/09813453A

Patent No. US20020168681A1
GENERAL INFORMATION:
APPLICANT: Yocum, R. Rogers
TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
FILE REFERENCE: OGZ-001
CURRENT APPLICATION NUMBER: US/09/813,453A
PRIOR FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: US 60/227,860
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 09/667,569
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 51
LENGTH: 260
TYPE: PRT
ORGANISM: Caulobacter crescentus
US-09-813-453A-51

Query Match 38.5%; Score 517; DB 9; Length 260;
Best Local Similarity 42.4%; Pred. No. 1.6e-37;
Matches 112; Conservative 50; Mismatches 90; Indels 12; Gaps 5;
QY 1 MLLTIDVGNTHVTLGLFDGEDIVEHWRISTDSRRRTADELAVLLQGLMGHPLGLGDELGDG 59
Db 1 MLLATBQGNNTMTFAIHGASWVAQWRSATSTRTADEYVWLSQLSLM-----QGLGFR 55
QY 60 GDGTAICATVPSVLHRELVTRRYGDPVAVLVEPCVGTGVPILTDHPKEVGADRLINAV 119
Db 56 AIDAVIISVVSQSIENLRLNLSRYNVEPLVIGE--NAKLGIDVRIEKPSEAGADRLVNA 114
QY 120 VAVELYGGPAIVDFGTATTDAVSARGEYIGGVAPLVEPCVGTGVPILTDHPKEVGADRLINAV 179
Db 115 IGAANYVPGVPLVIDSSTATTDDIVADGAFEGGIIAPGLINSMQALHEAAKLPRIATIQ 174
QY 180 RP--RSVIGKNTVAMQSGIVYGFAGQVDGVNRMARELADDPDDVTVIATGGLAPMYLG 237
Db 175 RPAGNRIVGTDTVSAMQSGVFWGYSILIEGLV---ARIKAERGEPTVIATGTVASLFEG 231
QY 238 ESSVIDEHEPWLTLMLGLRLVYERN 261
Db 232 ATDSIDHFDSDLTIRGLEIYRNN 255

Query Match 37.1%; Score 498.5; DB 9; Length 250;
Best Local Similarity 43.3%; Pred. No. 6.2e-36;
Matches 104; Conservative 50; Mismatches 73; Indels 13; Gaps 4;
US-09-813-453A-3
; Sequence 3, Application US/09813453A
Patent No. US20020168681A1
GENERAL INFORMATION:
APPLICANT: Yocum, R. Rogers
TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
FILE REFERENCE: OGZ-001
CURRENT APPLICATION NUMBER: US/09/813,453A
PRIOR FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: US 60/227,860
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 09/667,569
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 250
TYPE: PRT
ORGANISM: Clostridium acetobutylicum
US-09-813-453A-3

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

!OM protein - protein search, using sw model

Run on: June 24, 2003, 21:46:39 ; Search time 29.4306 Seconds
(without alignments)
1177.181 Million cell updates/sec

Title: US-09-813-453A-51
Perfect score: 1299
Sequence: 1 MLLAIQGNNTMFAIHGDA.....SDLTIRGLLEIYRNTIAES 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1299	100.0	260	AAU91173	Pantothenate kinase
2	658	50.7	258	AAU91153	Rhodobacter capsul
3	598	46.0	254	AAU91171	Pantothenate kinase
4	577	44.4	258	AAU01243	B. subtilis novel
5	577	44.4	258	AAU91149	Bacillus subtilis
6	575	44.3	258	AAU91172	Pantothenate kinase
7	566	43.6	262	AAU91170	Pantothenate kinase
8	561	43.2	255	AAU91154	Geobacter sulfurre
9	523.5	40.3	256	AAU91175	Pantothenate kinase
10	523	40.3	259	ABB47661	Listeria monocytog.

11	517	39.8	265	23	AAU91151	Streptomyces coeli
12	492	37.9	233	23	AAU91163	Pantothenate kinase
13	478	36.8	219	23	AAU91176	Pantothenate kinase
14	460.5	35.5	250	23	AAU91150	Clostridium acetob
15	396.5	30.5	262	23	AAU91155	Deinococcus radiop
16	382.5	29.4	272	23	AAU91152	Mycobacterium tube
17	378.5	29.1	272	22	AAU91125	Mycobacterium tube
18	359.5	27.7	246	23	AAU91156	Thermotoga maritim
19	337.5	26.0	273	23	AAU91157	Treponema pallidum
20	307.5	23.7	212	23	AAU91177	Pantothenate kinase
21	261	20.1	257	23	AAU91174	Pantothenate kinase
22	227	17.5	241	23	AAU91179	Pantothenate kinase
23	198	15.2	244	23	AAU91168	Pantothenate kinase
24	189.5	14.6	262	23	AAU91158	Borrelia burgdorfe
25	185	14.2	455	20	AAU91157	Neisseria gonorrhoe
26	185	14.2	455	21	AAU91158	Neisseria gonorrhoe
27	185	14.2	460	23	AAU91167	Pantothenate kinase
28	185	14.2	592	20	AAU91168	Neisseria gonorrhoe
29	185	14.2	592	21	AAU91169	Neisseria gonorrhoe
30	175	13.5	455	21	AAU91170	Neisseria meningit
31	175	13.5	592	20	AAU91171	Neisseria meningit
32	175	13.5	592	21	AAU91172	Neisseria meningit
33	175	13.5	592	23	AAU91166	Pantothenate kinase
34	173.5	13.4	389	21	AAU91167	Neisseria meningit
35	172	13.2	592	20	AAU91168	Neisseria meningit
36	172	13.2	592	21	AAU91169	Neisseria meningit
37	172	13.2	592	23	AAU91169	Pantothenate kinase
38	160	12.3	248	23	AAU91164	Pantothenate kinase
39	153.5	11.8	249	23	AAU91178	Pantothenate kinase
40	147	11.3	257	23	AAU91170	Synechocystis pant
41	141	10.9	249	23	AAU91182	Pantothenate kinase
42	140.5	10.8	242	23	AAU91180	Pantothenate kinase
43	134.5	10.4	267	23	AAU91162	Bordetella pertussis
44	125.5	9.7	229	23	AAU91159	Aquifex aeolicus p
45	116.5	9.0	189	20	AAU91164	Neisseria meningit

ALIGNMENTS

RESULT 1
AAU91173
ID AAU91173 standard; Protein; 260 AA.
AC AAU91173;
XX
XX 05-JUN-2002 (first entry)
DT
XX
XX Pantothenate kinase (Coax) #11.
DE
XX
XX Pantothenate kinase; Coax; antibiotic; antimicrobial;
KW Pantothenate kinase modulator; coenzyme A; bactericidal compound.
XX
XX Caulobacter crescentus.
OS
XX
XX WO200216601-A2.
PN
XX
XX 28-FEB-2002.
PD
XX
XX 24-AUG-2001; 2001WO-US26531.
PF
XX
XX 24-AUG-2000; 2000US-227860P.
PR 20-MAR-2001; 2001US-0813453.
XX
XX (OMNI-) OMNIGENE BIOPRODUCTS INC.
PA
XX
XX Yocum RR, Patterson TA;
PI
XX
XX WPI; 2002-269358/31.
DR N-PSDB; ABK54194.
XX
XX Identifying potential antibiotic or antimicrobial agent, comprises
PT contacting composition comprising pantothenate kinase (Coax) protein

PT with test compound and identifying inhibitor of the Coax protein -
 XX Claim 10; Page 102-103; 128pp; English.
 XX The invention describes assays for identifying a (potential) antibiotic
 CC comprising contacting an assay composition comprising a pantothenate
 CC kinase (Coax) protein with a test compound, and determining the ability
 CC of the test compound to inhibit the activity of the Coax protein, an
 CC essential enzyme for the production of coenzyme A. Coax protein is a
 CC valuable target for identifying bactericidal compounds. Coax modulating
 CC agents can be used in an infectious animal model to determine the
 CC efficacy, toxicity, or side effects of treatment with such an agent. This
 CC is the amino acid sequence of a pantothenate kinase (Coax) protein
 CC described in the invention.

SQ Sequence 260 AA;
 Query Match 100.0%; Score 1299; DB 23; Length 260;
 Best Local Similarity 100.0%; Pred. No. 8.3e-133; Mismatches 0; Indels 0; Gaps 0;
 Matches 260; Conservative 0;
 Db 1 MLLAEQNTNTMFAIHDCGASVQWRSATSTRTADEYVWLSQGLGFRDAIDAV 60
 1 MLLAEQNTNTMFAIHDCGASVQWRSATSTRTADEYVWLSQGLGFRDAIDAV 60
 Qy 61 IISVVPOSIFNRLNLSRRYFNVEPLVIGENAKLIDVRIKPSSEAGADRLVNAIGAAMV 120
 Db 61 IISVVPOSIFNRLNLSRRYFNVEPLVIGENAKLIDVRIKPSSEAGADRLVNAIGAAMV 120
 Qy 121 YPGPLVWIDSGTATTFDVAADGAFEGGIIAPGINSQALHEAAKLPRIAIORPAGNR 180
 Db 121 YPGPLVWIDSGTATTFDVAADGAFEGGIIAPGINSQALHEAAKLPRIAIORPAGNR 180
 Qy 181 IVGTDVTSAMOSGVFWGYSISLIEGLVARIKAEGERPMTVIATGGVASFEGATDSIDHFD 240
 Db 181 IVGTDVTSAMOSGVFWGYSISLIEGLVARIKAEGERPMTVIATGGVASFEGATDSIDHFD 240
 Qy 241 SDLTIRGLLEIYRRNTIAES 260
 Db 241 SDLTIRGLLEIYRRNTIAES 260

RESULT 2
 AAU91153
 ID AAU91153 standard; Protein; 258 AA.
 XX AAU91153;
 XX 05-JUN-2002 (first entry)
 Rhodobacter capsulatus pantothenate kinase Coax.
 XX Pantothenate kinase; Coax; antibiotic; antimicrobial;
 KW Pantothenate kinase modulator; coenzyme A; bactericidal compound.
 XX Rhodobacter capsulatus.
 OS WO200216601-A2.
 PN 28-FEB-2002.
 PD 24-AUG-2001; 2001WO-US26531.
 PF 24-AUG-2000; 2000US-227860P.
 PR 20-MAR-2001; 2001US-0813453.
 XX (OMNI-) OMNIGENE BIOPRODUCTS INC.
 PA Yocum RR, Patterson TA;
 PI WPI; 2002-269358/31.
 XX Identifying potential antibiotic or antimicrobial agent, comprises

PT contacting composition comprising pantothenate kinase (Coax) protein
 PT with test compound and identifying inhibitor of the Coax protein -
 XX Claim 10; Page 71-72; 128pp; English.
 XX The invention describes assays for identifying a (potential) antibiotic
 CC comprising contacting an assay composition comprising a pantothenate
 CC kinase (Coax) protein with a test compound, and determining the ability
 CC of the test compound to inhibit the activity of the Coax protein, an
 CC essential enzyme for the production of coenzyme A. Coax protein is a
 CC valuable target for identifying bactericidal compounds. Coax modulating
 CC agents can be used in an infectious animal model to determine the
 CC efficacy, toxicity, or side effects of treatment with such an agent. This
 CC is the amino acid sequence of a pantothenate kinase (Coax) protein
 CC described in the invention.

SQ Sequence 258 AA;
 Query Match 50.7%; Score 658; DB 23; Length 258;
 Best Local Similarity 50.8%; Pred. No. 4.9e-63; Mismatches 48; Indels 4; Gaps 3;
 Matches 130; Conservative 48;
 Qy 1 MLLAEQNTNTMFAIHDCGASVQWRSATSTRTADEYVWLSQGLGFRDAIDAV 60
 Db 1 MLLAEQNTNTMFAIHDCGASVQWRSATSTRTADEYVWLSQGLGFRDAIDAV 60
 Qy 61 IISVVPOSIFNRLNLSRRYFNVEPLVIGENAKLIDVRIKPSSEAGADRLVNAIGAAM 119
 Db 61 IISVVPOSIFNRLNLSRRYFNVEPLVIGENAKLIDVRIKPSSEAGADRLVNAIGAAM 119
 Qy 120 YPGPLVWIDSGTATTFDVAADGAFEGGIIAPGINSQALHEAAKLPRIAIORPAGN 179
 Db 120 YPGPLVWIDSGTATTFDVAADGAFEGGIIAPGINSQALHEAAKLPRIAIORPAGN 179
 Qy 180 RVGTDVTSAMOSGVFWGYSISLIEGLVARIKAEGERPMTVIATGGVASFEGATDSIDHF 239
 Db 180 RVGTDVTSAMOSGVFWGYSISLIEGLVARIKAEGERPMTVIATGGVASFEGATDSIDHF 239
 Qy 240 SDLTIRGLLEIYRRN 255
 Db 238 EDDLTMHGLRLIFDYN 253

RESULT 3
 AAU91171
 ID AAU91171 standard; Protein; 254 AA.
 XX AAU91171;
 XX 05-JUN-2002 (first entry)
 Pantothenate kinase (Coax) #9.
 XX Pantothenate kinase; Coax; antibiotic; antimicrobial;
 KW Pantothenate kinase modulator; coenzyme A; bactericidal compound.
 XX Bacillus halodurans.
 OS WO200216601-A2.
 PN 28-FEB-2002.
 PD 24-AUG-2001; 2001WO-US26531.
 PF 24-AUG-2000; 2000US-227860P.
 PR 20-MAR-2001; 2001US-0813453.
 XX (OMNI-) OMNIGENE BIOPRODUCTS INC.
 PA Yocum RR, Patterson TA;
 PI WPI; 2002-269358/31.
 XX N-PSDB; ABK54192.

[illegible]

PR 20-MAR-2001; 2001US-0813453.
XX (OMNI-) OMNIGENE BIOPRODUCTS INC.
XX Yocum RR, Patterson TA;
PI WPI: 2002-269358/31.
XX N-PSDB; ABK54168.
DR Identifying potential antibiotic or antimicrobial agent, comprises
XX contacting composition comprising pantothenate kinase (Coax) protein
PT with test compound and identifying inhibitor of the Coax protein
XX
PS Claim 10; Page 67-68; 128pp; English.
XX
CC The invention describes assays for identifying a (potential) antibiotic
CC comprising contacting an assay composition comprising a pantothenate
CC kinase (Coax) protein with a test compound, and determining the ability
CC of the test compound to inhibit the activity of the Coax protein, an
CC essential enzyme for the production of coenzyme A. Coax protein is a
CC valuable target for identifying bactericidal compounds. Coax modulating
CC agents can be used in an infectious animal model to determine the
CC efficacy, toxicity, or side effects of treatment with such an agent. This
CC is the amino acid sequence of a pantothenate kinase (Coax) protein
CC described in the invention.
XX
SQ Sequence 258 AA;
Query Match 44.4%; Score 577; DB 23; Length 258;
Best Local Similarity 44.4%; Pred. No. 3.2e-54;
Matches 114; Conservative 50; Mismatches 91; Indels 2; Gaps 1;
QY 1 MLLAIEQGNNTWFAIHGDSWVAQWRSATSTRTADEYVYVWLSQGLGFRAIDAV 60
DB 1 LLLLVIVGNTNTVLGYVHDGKLEYHWRIETSRHKTEDEFGMLILSLFDHSGLMFEQIDGI 60
QY 61 IISVVVQSFIFNLNLSRRYFNVEPLVIGENAKLGIDVRIEKPSEAGADRLVNAIGAAV 120
DB 61 IISVVVPPIMFALERMCTFYFIEPQVIGPMKTLGNIYDNPKEVGADRVNAVAIHL 120
QY 121 YPGPLWVDSGTATTFDVAADGAFEGGIIAPGINSQALHEAAKLPRIATQRPAGNR 180
DB 121 YGNPLIVDFGATTCYIDENKQYMGGAIAFGITISTEALYSRAAKLPRIETRP -DN 178
QY 181 IVGTDVTSAMQSGVFWGYISLIEGLVARIKAERGEPMVTIATGGVASFEGATDSIDHFD 240
DB 179 IIGKNTVSAMQSGILFYGVQVEGIVKRMQAKQDKLVATGGLAPLANESDCIDIVD 238
QY 241 SLDLTIRGLLEIYRRNTI 257
239 PFLTGLKLEIYERNRV 255
RESULT 6
AAU91172
ID AAU91172 standard; Protein; 258 AA.
XX AAU91172;
AC
XX
DT 05-JUN-2002 (first entry)
XX
DE Pantothenate kinase (Coax) #10.
XX Pantothenate kinase; Coax; antibiotic; antimicrobial;
KW pantothenate kinase modulator; coenzyme A; bactericidal compound.
XX
OS Bacillus stearothermophilus.
XX
PN WO200216601-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26531.

XX 24-AUG-2000; 2000US-227860P.
PR 20-MAR-2001; 2001US-0813453.
XX
XX (OMNI-) OMNIGENE BIOPRODUCTS INC.
XX Yocum RR, Patterson TA;
PI WPI: 2002-269358/31.
XX N-PSDB; ABK54193.
DR Identifying potential antibiotic or antimicrobial agent, comprises
XX contacting composition comprising pantothenate kinase (Coax) protein
PT with test compound and identifying inhibitor of the Coax protein
XX
PS Claim 10; Page 101-102; 128pp; English.
XX
CC The invention describes assays for identifying a (potential) antibiotic
CC comprising contacting an assay composition comprising a pantothenate
CC kinase (Coax) protein with a test compound, and determining the ability
CC of the test compound to inhibit the activity of the Coax protein, an
CC essential enzyme for the production of coenzyme A. Coax protein is a
CC valuable target for identifying bactericidal compounds. Coax modulating
CC agents can be used in an infectious animal model to determine the
CC efficacy, toxicity, or side effects of treatment with such an agent. This
CC is the amino acid sequence of a pantothenate kinase (Coax) protein
CC described in the invention.
XX
SQ Sequence 258 AA;
Query Match 44.3%; Score 575; DB 23; Length 258;
Best Local Similarity 44.1%; Pred. No. 5.3e-54;
Matches 113; Conservative 52; Mismatches 89; Indels 2; Gaps 1;
QY 1 MLLAIEQGNNTWFAIHGDSWVAQWRSATSTRTADEYVYVWLSQGLGFRAIDAV 60
DB 1 MIFVLVGVNTNTVLGYVHDGDELKHHWRIETSRKTEDEYGMIMKALLNHVGLQFSDIRGI 60
QY 61 IISVVVQSFIFNLNLSRRYFNVEPLVIGENAKLGIDVRIEKPSEAGADRLVNAIGAAV 120
DB 61 IISVVVPPIMFALERMCTFYFIEPQVIGPMKTLGNIYDNPKEVGADRVNAVAIHL 120
QY 121 YPGPLWVDSGTATTFDVAADGAFEGGIIAPGINSQALHEAAKLPRIATQRPAGNR 180
DB 121 YGSLIIVDFGTATTCYIDENKQYMGGAIAFGITISTEALYSRAAKLPRIETARP --DD 178
QY 181 IVGTDVTSAMQSGVFWGYISLIEGLVARIKAERGEPMVTIATGGVASFEGATDSIDHFD 240
DB 179 IIGKNTVSAMQAGILFYGVQVEGIVSRMKAKSKIPPKVATGGLAPLIASESDIIVVD 238
QY 241 SLDLTIRGLLEIYRRNT 256
239 PFLTGLKLEIYKNT 254
RESULT 7
AAU91170
ID AAU91170 standard; Protein; 262 AA.
XX AAU91170;
AC
XX
DT 05-JUN-2002 (first entry)
XX
DE Pantothenate kinase (Coax) #8.
XX Pantothenate kinase; Coax; antibiotic; antimicrobial;
KW pantothenate kinase modulator; coenzyme A; bactericidal compound.
XX
OS Bacillus anthracis.
XX
PN WO200216601-A2.
XX
PD 28-FEB-2002.

XX PF 24-AUG-2001; 2001WO-US26531.
XX PR 24-AUG-2000; 2000US-227860P.
XX PR 20-MAR-2001; 2001US-0813453.
XX PA (OMNI-) OMNIGENE BIOPRODUCTS INC.
XX PI Yocum RR, Patterson TA;
XX DR WPI; 2002-269358/31.
XX DR N-PSDB; ABK54191.
XX PT Identifying potential antibiotic or antimicrobial agent, comprises
XX PT contacting composition comprising pantothenate kinase (Coax) protein
XX PT with test compound and identifying inhibitor of the Coax protein
XX PS Claim 8; Page 98-99; 128pp; English.
XX C The invention describes assays for identifying a (potential) antibiotic
XX C comprising contacting an assay composition comprising a pantothenate
XX C kinase (Coax) protein with a test compound, and determining the ability
XX C of the test compound to inhibit the activity of the Coax protein, an
XX C essential enzyme for the production of coenzyme A. Coax protein is a
XX C valuable target for identifying bactericidal compounds. Coax modulating
XX C agents can be used in an infectious animal model to determine the
XX C efficacy, toxicity, or side effects of treatment with such an agent. This
XX C is the amino acid sequence of a pantothenate kinase (Coax) protein
XX C described in the invention.
XX C Sequence 262 AA;
Query Match 43.6%; Score 566; DB 23; Length 262;
Best Local Similarity 42.0%; Pred. No. 5.2e-53;
Matches 107; Conservative 55; Mismatches 91; Indels 2; Gaps 1;
QY 1 MLLAIEQNTNMFALHDGASWVAQWRSATSTRTADEYVWVLSQGLGFRAIDAV 60
DB 1 MLFVLDVGNNAVGVFEGELRQHWRMETDRHKTEDEYGMVQLLEHGLSFEDVKGI 60
QY 61 IISVVPPQSFIFNLNLSRRYFNVEPLVIGENAKLGIDVRIEKPSEAGADRLVNAIGAAMV 120
DB 61 IVSSVPPIMFALERMCERYFKPLVGVGFKTGKLNKIKYENPREVGADRLVNAVAGIHL 120
QY 121 YGPLVVDISGATTDFDVAADGAFEGGIIAPGINLSMQALHAAKLPRIATIORPAGNR 180
DB 121 YGSLIIVDFGATTATTCYINEEKHYMGVITPGIMISAEALYSRAAKLPRIETKPS-- 178
QY 181 IVGTDTVSAMQSGVFWGYSISLIEGLVARIKAEERGEPTMTVIATGGVASLFEGATDSIDHFD 240
DB 179 VVGKNTVSAMQSGILYGYGVQVEGIVKRMKEAKQEPKVIATGGLAKLISESNVDVVD 238
QY 241 SDLTIRGLLEIYRN 255
DB 239 PFLTKGLYLYERN 253
RESULT 8
AAU91154
ID AAU91154 standard; Protein; 255 AA.
XX AC AAU91154;
XX DT 05-JUN-2002 (first entry)
XX DE Geobacter sulfurreducens pantothenate kinase Coax.
XX KW Pantothenate kinase; Coax; antibiotic; antimicrobial;
XX KW pantothenate kinase modulator; coenzyme A; bactericidal compound.
XX OS Geobacter sulfurreducens.
XX PN WO200216601-A2.

XX PD 28-FEB-2002.
XX PF 24-AUG-2001; 2001WO-US26531.
XX PR 24-AUG-2000; 2000US-227860P.
XX PR 20-MAR-2001; 2001US-0813453.
XX PA (OMNI-) OMNIGENE BIOPRODUCTS INC.
XX PI Yocum RR, Patterson TA;
XX DR WPI; 2002-269358/31.
XX PT Identifying potential antibiotic or antimicrobial agent, comprises
XX PT contacting composition comprising pantothenate kinase (Coax) protein
XX PT with test compound and identifying inhibitor of the Coax protein
XX PS Claim 10; Page 72-73; 128pp; English.
XX C The invention describes assays for identifying a (potential) antibiotic
XX C comprising contacting an assay composition comprising a pantothenate
XX C kinase (Coax) protein with a test compound, and determining the ability
XX C of the test compound to inhibit the activity of the Coax protein, an
XX C essential enzyme for the production of coenzyme A. Coax protein is a
XX C valuable target for identifying bactericidal compounds. Coax modulating
XX C agents can be used in an infectious animal model to determine the
XX C efficacy, toxicity, or side effects of treatment with such an agent. This
XX C is the amino acid sequence of a pantothenate kinase (Coax) protein
XX C described in the invention.
XX C Sequence 255 AA;
Query Match 43.2%; Score 561; DB 23; Length 255;
Best Local Similarity 44.7%; Pred. No. 1.8e-52;
Matches 114; Conservative 50; Mismatches 89; Indels 2; Gaps 1;
QY 1 MLLAIEQNTNMFALHDGASWVAQWRSATSTRTADEYVWVLSQGLGFRAIDAV 60
DB 1 MLLVIDVGNNAVGVFEGELRQHWRMETDRHKTEDEYGMVQLLEHGLSFEDVKGI 60
QY 61 IISVVPPQSFIFNLNLSRRYFNVEPLVIGENAKLGIDVRIEKPSEAGADRLVNAIGAAMV 120
DB 61 IISVVPPQSFIFNLNLSRRYFNVEPLVIGENAKLGIDVRIEKPSEAGADRLVNAIGAAMV 120
QY 121 YGPLVVDISGATTDFDVAADGAFEGGIIAPGINLSMQALHAAKLPRIATIORPAGNR 180
DB 121 YGSLIIVDFGATTATTCYINEEKHYMGVITPGIMISAEALYSRAAKLPRIETKPS-- 178
QY 181 IVGTDTVSAMQSGVFWGYSISLIEGLVARIKAEERGEPTMTVIATGGVASLFEGATDSIDHFD 240
DB 179 IIAINTVNSMQAGIYGYGVQVEGIVKRMKEAKQEPKVIATGGLAKLISESNVDVVD 238
QY 241 SDLTIRGLLEIYRN 255
DB 239 EYLTLEGLIYERN 253
RESULT 9
AAU91175
ID AAU91175 standard; Protein; 256 AA.
XX AC AAU91175;
XX DT 05-JUN-2002 (first entry)
XX DE Pantothenate kinase (Coax) #13.
XX KW Pantothenate kinase; Coax; antibiotic; antimicrobial;
XX KW pantothenate kinase modulator; coenzyme A; bactericidal compound.
XX OS Clostridium difficile.
XX PN

PN WO200216601-A2.
XX 28-FEB-2002.
XX 24-AUG-2001; 2001WO-US26531.
XX 24-AUG-2000; 2000US-227860P.
XX 20-MAR-2001; 2001US-0813453.
XX (OMNI-) OMNIGENE BIOPRODUCTS INC.
XX Yocum RR, Patterson TA;
XX WPI; 2002-269358/31.
XX N-PSDB; ABK54196.
XX Identifying potential antibiotic or antimicrobial agent, comprises
XX contacting composition comprising pantothenate kinase (Coax) protein
XX with test compound and identifying inhibitor of the Coax protein -
XX Claim 6; Page 105; 128pp; English.
XX The invention describes assays for identifying a (potential) antibiotic
XX comprising contacting an assay composition comprising a pantothenate
XX kinase (Coax) protein with a test compound, and determining the ability
XX of the test compound to inhibit the activity of the Coax protein, an
XX essential enzyme for the production of coenzyme A. Coax protein is a
XX valuable target for identifying bactericidal compounds. Coax modulating
XX agents can be used in an infectious animal model to determine the
XX efficacy, toxicity, or side effects of treatment with such an agent. This
XX is the amino acid sequence of a pantothenate kinase (Coax) protein
XX described in the invention.
XX Sequence 256 AA;
XX Query Match 40.3%; Score 523.5; DB 23; Length 256;
XX Best Local Similarity 40.6%; Pred. No. 2.1e-48;
XX Matches 104; Conservative 53; Mismatches 96; Indels 3; Gaps 2;
QY 1 MLLATEQNTMTFAIHDCASWVAQWRSATSTRDAEYVWLSQLSGGLGFRAIDAV 60
DB 1 MLLVDFVGNCTNVLGIYKDKLVNRYRIKTDREKTSDEYGLISNLFYDYNISDIDDV 60
QY 61 IISVVVPOSIFNLRLNLSRYFVNEPLVIGENAKLGDVRIEKPSPAGADRLVNAIGAAMV 120
DB 61 IISVVVPMHSLNFCIKYCKQPLVIGPGIKTGLNPKYDNPQVGDRIYVNAVAGIEK 120
QY 121 YPGPLVWIDSGTATTFDVAADGAFEGGIIAPGINSMLQALHEAAAKLPRIAIQRPAGNR 180
DB 121 YGAPSLVDFGTATTFCAISEKGYLGTTIAPGINKISSEALFQSAKLPRLVELAKP--GM 178
QY 181 IVGTDTVSAMOSGVFWGYISLIEGLVARIKAERG-EPMTVIATGGVASFEGATSDIDHF 239
DB 179 TICKSTVSAMOSGIYGVGLVDLKIISIMKLNLCDDVKVYATGGLAKLIASETSIDYV 238
QY 240 DSDLTIRGLLEIYRN 255
DB 239 DGFLTLEGLRIYEKN 254
RESULT 10
ABB47661
ID ABB47661 standard; Protein; 259 AA.
XX ABB47661;
XX 05-FEB-2002 (first entry)
XX Listeria monocytogenes protein #365.
XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX vitamin B12; bacterial infection; disease.
XX

OS Listeria monocytogenes.
XX WO200177335-A2.
XX 18-OCT-2001.
XX 11-APR-2001; 2001WO-FR01118.
XX 11-APR-2000; 2000FR-0004629.
XX (INSP) INST PASTEUR.
XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
XX Rose M, Voss H;
XX WPI; 2002-010914/01.
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX and prevention of Listeria and related bacterial infections, and
XX related polypeptides -
XX Claim 6; SEQ ID No 366; 192pp; French.
XX The present invention relates to the genome sequence of Listeria
XX monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
XX it are useful for selecting probes and primers for detecting genes in L.
XX monocytogenes and related organisms, and for studying genetic
XX polymorphisms and other genomes. The present sequence is a protein
XX encoded by the genome sequence of the present invention. Proteins
XX expressed from the genome sequence are useful for raising specific
XX antibodies, identification of L. monocytogenes and related organisms, and
XX for biosynthesis and biodegradation, especially biosynthesis of Vitamin
XX #12. The genome sequence and proteins encoded by it are also useful for
XX selecting compounds that regulate gene expression and cell replication
XX and modulate L. monocytogenes-related diseases. In addition, the genome
XX sequence and proteins encoded by it are useful in pharmaceutical and
XX vaccine compositions for the treatment or prevention of infections by L.
XX monocytogenes and related organisms.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/published_pct_sequences.
XX Sequence 259 AA;
XX Query Match 40.3%; Score 523; DB 23; Length 259;
XX Best Local Similarity 39.2%; Pred. No. 2.5e-48;
XX Matches 100; Conservative 59; Mismatches 94; Indels 2; Gaps 1;
QY 1 MLLATEQNTMTFAIHDCASWVAQWRSATSTRDAEYVWLSQLSGGLGFRAIDAV 60
DB 1 MLLVDFVGNCTNVLGIYKDKLVNRYRIKTDREKTSDEYGLISNLFYDYNISDIDDV 60
QY 61 IISVVVPOSIFNLRLNLSRYFVNEPLVIGENAKLGDVRIEKPSEAGADRLVNAIGAAMV 120
DB 61 IISVVVPMHMETMCMVRYFNIRPLVIGIKTGLNKLKVDNPREIGSDRIVNAASEE 120
QY 121 YPGPLVWIDSGTATTFDVAADGAFEGGIIAPGINSMLQALHEAAAKLPRIAIQRPAGNR 180
DB 121 YGTPVIVDFGTATTFICYIDESVYOGGAIFGIMISTEALYNRAKLPDVDAE--SSQ 178
QY 181 IVGTDTVSAMOSGVFWGYISLIEGLVARIKAERGEPMTVIATGGVASFEGATSDIDHF 240
DB 179 IICKSTVSSMQAGIFYGFVQCEGIIAEMKQSNASPVVAVTGTGLARMITEKSSAVDILD 238
QY 241 SDLTIRGLLEIYRN 255
DB 239 PFTLTKGLELLYRN 253

00Z VVVVVVAT06EAPOA10JTTCSCB'WBCA INVDOT 6/T 000

RESULT 13

AAU91176
ID AAU91176 standard; Protein; 219 AA.

XX AC AAU91176;
XX DT 05-JUN-2002 (first entry)

XX DE Pantothenate kinase (Coax) #14.
XX KW Pantothenate kinase; Coax; antibiotic; antimicrobial;
XX KW Pantothenate kinase modulator; coenzyme A; bactericidal compound.

XX OS Dehalococcoides ethenogenes.
XX PN WO200216601-A2.
XX PD 28-FEB-2002.
XX PR 24-AUG-2001; 2001WO-US26531.
XX PR 24-AUG-2000; 2000US-227860P.
XX PR 20-MAR-2001; 2001US-0813453.
XX PA (OMNI-) OMNIGENE BIOPRODUCTS INC.
XX PI Yocum RR, Patterson TA;
XX DR WPI; 2002-269358/31.
XX DR N-PSDB; ABK54197.

XX PT Identifying potential antibiotic or antimicrobial agent, comprises
XX PT contacting composition comprising pantothenate kinase (Coax) protein
XX PT with test compound and identifying inhibitor of the Coax protein -
XX PS Claim 10; Page 106-107; 128pp; English.

XX CC The invention describes assays for identifying a (potential) antibiotic
XX CC comprising contacting an assay composition comprising a pantothenate
XX CC kinase (Coax) protein with a test compound, and determining the ability
XX CC of the test compound to inhibit the activity of the Coax protein, an
XX CC essential enzyme for the production of coenzyme A. Coax protein is a
XX CC valuable target for identifying bactericidal compounds. Coax modulating
XX CC agents can be used in an infectious animal model to determine the
XX CC efficacy, toxicity, or side effects of treatment with such an agent. This
XX CC is the amino acid sequence of a pantothenate kinase (Coax) protein
XX CC described in the invention.

Sequence 219 AA;

Query Match 36.8%; Score 478; DB 23; Length 219;

Best Local Similarity 44.9%; Pred. No. 1.5e-43;

Matches 96; Conservative 41; Mismatches 75; Indels 2; Gaps 1;

QY 2 LLALTEOGTNTMFAHDGASVWAQWRSATSTRTADEYVWLSQLLSMOGLGFRDAVI 61

Db 5 LVAVDIGNTSNIGIFEGEKLNNHLSGVAQWMADEYASLLGLLQHAGIHPEELNRYI 64

QY 62 ISSVVPQISIFNLNLSRRYFNVEPLVIGENAKLGIDVRIEKPSEAGADRLVNAIGAAMY 121

Db 65 MCSVVPPLTTTFEEFKSYFAALVLVVGAGIKSGVKVRMDNPREVGADRIVNAAGRVLY 124

QY 122 PGPVLVWIDSGPATTFTDVAADGAFEGGIAPGINLSMOALHEAAKLPRIATQRPAGNRI 181

Db 125 PGACIIIVDMGPTATFTDLSSEGAYIGGAIAPGIATSAQIAEKTSLKPIELIRPA--KV 182

QY 182 VGTDTVSAMQSGVFWGYISLIEGLVARIKAE 215

Db 183 IGSNTVSAMQSGIYFVGLVLEELVRRIQTQLGQ 216

RESULT 14

AAU91150

ID AAU91150 standard; Protein; 250 AA.

XX AC AAU91150;
XX DT 05-JUN-2002 (first entry)

XX DE Clostridium acetobutylicum pantothenate kinase Coax.
XX KW Pantothenate kinase; Coax; antibiotic; antimicrobial;
XX KW Pantothenate kinase modulator; coenzyme A; bactericidal compound.

XX OS Clostridium acetobutylicum.
XX PN WO200216601-A2.
XX PD 28-FEB-2002.
XX PR 24-AUG-2001; 2001WO-US26531.
XX PR 24-AUG-2000; 2000US-227860P.
XX PR 20-MAR-2001; 2001US-0813453.
XX PA (OMNI-) OMNIGENE BIOPRODUCTS INC.
XX PI Yocum RR, Patterson TA;
XX DR WPI; 2002-269358/31.

XX PT Identifying potential antibiotic or antimicrobial agent, comprises
XX PT contacting composition comprising pantothenate kinase (Coax) protein
XX PT with test compound and identifying inhibitor of the Coax protein -
XX PS Claim 10; Page 68-69; 128pp; English.

XX CC The invention describes assays for identifying a (potential) antibiotic
XX CC comprising contacting an assay composition comprising a pantothenate
XX CC kinase (Coax) protein with a test compound, and determining the ability
XX CC of the test compound to inhibit the activity of the Coax protein, an
XX CC essential enzyme for the production of coenzyme A. Coax protein is a
XX CC valuable target for identifying bactericidal compounds. Coax modulating
XX CC agents can be used in an infectious animal model to determine the
XX CC efficacy, toxicity, or side effects of treatment with such an agent. This
XX CC is the amino acid sequence of a pantothenate kinase (Coax) protein
XX CC described in the invention.

Sequence 250 AA;

Query Match 35.5%; Score 460.5; DB 23; Length 250;

Best Local Similarity 39.5%; Pred. No. 1.5e-41;

Matches 92; Conservative 53; Mismatches 81; Indels 7; Gaps 2;

QY 1 MLALTEOGTNTMFAHDGASVWAQWRSATSTRTADEYVWLSQLLSMOGLGFRDAVI 60

Db 18 VILVLDVGNINVLGIYNDTKLTAEWRLSTDVRSADYEYGIQVNLNFOQDKLDTLVEGV 77

QY 61 IISVVVPQISIFNLNLSRRYFNVEPLVIGENAKLGIDVRIEKPSEAGADRLVNAIGAAMY 120

Db 78 IISVVVPNIMYSLEHMIRKFKINPLVVGPGIKTGINIKYDNPKEVGADRIVNAVAHEI 137

QY 121 YPGPLVWIDSGPATTFTDVAADGAFEGGIAPGINLSMOALHEAAKLPRIATQRPAGNR 180

Db 138 YKRSLLIIDFGTATTCVAVRENGDYLGAICPGIKVSSSEALFEKAALPRVELIKPA--Y 195

QY 181 IVGTDTVSAMQSGVFWGYISLIEGLVARIK-----AERGPMTVIATGGVSL 228

Db 196 AICKNTIISQSIQYIVRYLRQVKYLFELKLENLPDGRTRTSLVLTATGGLAKL 248

RESULT 15

AAU91155
ID AAU91155 standard; Protein; 262 AA.

XX

AC AAU91155;
XX 05-JUN-2002 (first entry)
DT DE
XX DE
DE Deinococcus radiopugnans pantothenate kinase Coax.
XX
KW Pantothenate kinase; Coax; antibiotic; antimicrobial;
KW Pantothenate kinase modulator; coenzyme A; bactericidal compound.
XX
OS Deinococcus radiopugnans.
XX
PN WO200216601-A2.
XX
PD 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US26531.
XX
XX 24-AUG-2000; 2000US-227860P.
PR 20-MAR-2001; 2001US-0813453.
XX
X (OMNI-) OMNIGENE BIOPRODUCTS INC.
XA
XX Yocum RR, Patterson TA;
XX WPI; 2002-269358/31.
XX
PT Identifying potential antibiotic or antimicrobial agent, comprises
PT contacting composition comprising pantothenate kinase (Coax) protein
PT with test compound and identifying inhibitor of the Coax protein -
XX
PS Claim 10; Page 73-74; 128pp; English.
XX
CC The invention describes assays for identifying a (potential) antibiotic
CC comprising contacting an assay composition comprising a pantothenate
CC kinase (Coax) protein with a test compound, and determining the ability
CC of the test compound to inhibit the activity of the Coax protein, an
CC essential enzyme for the production of coenzyme A. Coax protein is a
CC valuable target for identifying bactericidal compounds. Coax modulating
CC agents can be used in an infectious animal model to determine the
CC efficacy, toxicity, or side effects of treatment with such an agent. This
CC is the amino acid sequence of a pantothenate kinase (Coax) protein
CC described in the invention.
XX
SQ Sequence 262 AA;

Query Match 30.5%; Score 396.5; DB 23; Length 262;
Best Local Similarity 36.1%; Pred. No. 1.5e-34;
Matches 92; Conservative 45; Mismatches 109; Indels 9; Gaps 5;

QY 2 LLAIEGNTNTMFAIHGASWVAQ-WRSATESTRTADEYVYVWLSQGLGFRIDAIV 60
DB III: III:
6 LLAVDIGNTTTVLGLADASCALHTWRTNREMLPDDLALQLHGLFTLAGAPIR-AA 63
QY 61 IISVVVPSQIFNLNLSRRYFNVPLVIGENAKLGIDVRIEKPSEAGADLVNAIGAAMV 120
DB :::::
64 VLSSVAPPVGENYALAKRHEMIDAFVSAENLPDVTVELDTPGSGADRLCNLFGAEK- 122
QY 121 YPGPL---VVIDSGTATTDIVAADGAFEGGIAPGINLSMOALHEAAKLPRIAIORPA 177
DB :
123 YLGGLDYAVVDFGTSTNDFVGRGRFFLGGLIATGAQVSADALFARAALPRITLQAP- 181
QY 178 GNRIVGTDVTSAMQSGVFWGYISLIEGLVARIKAEERGEPMVIATGVSASLFEGETSID 237
DB :
182 -ETAIGNVTHALQSLVFGYAEMVDGLLRIRRAELPGEAVAVATGGFSRTVGICQEID 240
QY 238 HFDSDLTIRGLLEIY 252
DB :::
241 YYDETTLRLGLVELW 255

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 21:27:26 ; Search time 27.207 Seconds
(without alignments)
1297.879 Million cell updates/sec

Title: US-09-813-453A-4
Perfect score: 1343
Sequence: 1 MLLTIDVGNTHVTLGFDGE.....EPWLTLMGLVLYERNVSRM 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
tal number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1343	100.0	265	23 AAU91151	Streptomyces coeli
2	698.5	52.0	272	23 AAU91152	Mycobacterium tube
3	694.5	51.7	272	22 AAG81225	Mycobacterium tube
4	663	49.4	258	23 AAU91172	Pantothenate kinase
5	662	49.3	254	23 AAU91171	Pantothenate kinase
6	655	48.8	258	22 AAU01243	B. subtilis novel
7	655	48.8	258	23 AAU91149	Bacillus subtilis
8	643.5	47.9	256	23 AAU91175	Pantothenate kinase
9	638	47.5	262	23 AAU91170	Pantothenate kinase
10	629	46.8	255	23 AAU91154	Geobacter sulfurre

11	596	44.4	259	23 AB47661	Listeria monocytog
12	554.5	41.3	233	23 AAU91163	Pantothenate kinase
13	527	39.2	238	23 AAU91153	Rhodobacter capsul
14	517	38.5	260	23 AAU91173	Pantothenate kinase
15	498.5	37.1	250	23 AAU91150	Clostridium acetob
16	492.5	36.7	219	23 AAU91176	Pantothenate kinase
17	409.5	30.5	262	23 AAU91155	Deinococcus radiop
18	342.5	25.5	246	23 AAU91156	Thermotoga maritim
19	295	22.0	237	23 AAU91174	Pantothenate kinase
20	291.5	21.7	273	23 AAU91157	Treponema pallidum
21	280	20.8	212	23 AAU91177	Pantothenate kinase
22	210	15.6	249	23 AAU91178	Pantothenate kinase
23	205.5	15.3	244	23 AAU91168	Pantothenate kinase
24	184	13.7	241	23 AAU91179	Pantothenate kinase
25	181.5	13.5	262	23 AAU91158	Borrelia burgdorfe
26	180.5	13.4	455	21 AAU74910	Neisseria meningit
27	180.5	13.4	592	20 AAU38616	Neisseria meningit
28	180.5	13.4	592	21 AAU74913	Neisseria meningit
29	180.5	13.4	592	23 AAU91166	Pantothenate kinase
30	179	13.3	455	20 AAU38617	Neisseria gonorrh
31	179	13.3	455	21 AAU74908	Neisseria gonorrh
32	179	13.3	460	23 AAU91167	Pantothenate kinase
33	179	13.3	592	20 AAU38618	Neisseria gonorrh
34	179	13.3	592	21 AAU74911	Neisseria gonorrh
35	177	13.2	592	20 AAU38615	Neisseria meningit
36	177	13.2	592	21 AAU74912	Neisseria meningit
37	177	13.2	592	23 AAU91169	Pantothenate kinase
38	176	13.1	249	23 AAU91182	Pantothenate kinase
39	174	13.0	257	23 AAU91160	Synechocystis pant
40	170.5	12.7	389	21 AAU74909	Neisseria meningit
41	139	10.3	242	23 AAU91180	Pantothenate kinase
42	138.5	10.3	248	23 AAU91184	Pantothenate kinase
43	136	10.1	189	20 AAU38614	Neisseria meningit
44	134	10.0	267	23 AAU91162	Bordetella pertussis
45	128.5	9.6	229	23 AAU91159	Aquifex aeolicus p

ALIGNMENTS

RESULT 1
AAU91151
ID AAU91151 standard; Protein; 265 AA.

XX AAU91151;

DT 05-JUN-2002 (first entry)

XX Streptomyces coelicolor pantothenate kinase Coax.

DE Pantothenate kinase; Coax; antibiotic; antimicrobial;

KW pantothenate kinase modulator; coenzyme A; bactericidal compound.

XX Streptomyces coelicolor.

PN WO200216601-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US26531.

XX 24-AUG-2000; 2000US-227860P.

XX 20-MAR-2001; 2001US-0813453.

XX (OMNI-) OMNIGENE BIOPRODUCTS INC.

XX Yocum RR, Patterson TA;

XX WPI; 2002-269358/31.

PT Identifying potential antibiotic or antimicrobial agent, comprises contacting composition comprising pantothenate kinase (Coax) protein with test compound and identifying inhibitor of the Coax protein -

```

XX PS Claim 10; Page 69-70; 128pp; English.
XX CC The invention describes assays for identifying a (potential) antibiotic
XX CC comprising contacting an assay composition comprising a pantothenate
XX CC kinase (Coax) protein with a test compound, and determining the ability
XX CC of the test compound to inhibit the activity of the Coax protein, an
XX CC essential enzyme for the production of coenzyme A. Coax protein is a
XX CC valuable target for identifying bactericidal compounds. Coax modulating
XX CC agents can be used in an infectious animal model to determine the
XX CC efficacy, toxicity, or side effects of treatment with such an agent. This
XX CC is the amino acid sequence of a pantothenate kinase (Coax) protein
XX CC described in the invention.
XX SQ Sequence 265 AA;

Query Match 100.0%; Score 1343; DB 23; Length 265;
Best Local Similarity 100.0%; Pred. No. 2.2e-122;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MLLTIDVGNTHVTLGLFDGDIVEHWRISTDSRRTADELAVLLOGLMGHPLLGDELGDG 60
1 MLLTIDVGNTHVTLGLFDGDIVEHWRISTDSRRTADELAVLLOGLMGHPLLGDELGDG 60
QY 61 IDGTAICATVPSVLHELREVTTRYGDPVAVLVEPGVKTGPILTDHPKVEGADRIINAV 120
121 AAVELYGGPAIVDFGTATTDDAVSARGEYIGGVIAPEGIEISVEALGVKGAQLRKIEVAR 180
Db 121 AAVELYGGPAIVDFGTATTDDAVSARGEYIGGVIAPEGIEISVEALGVKGAQLRKIEVAR 180
QY 181 PRSVIGNKTVEAMQSGIYVGFAGQVDGVVNRMAELADDDPDVTVIATGGLAPMVLGESS 240
181 PRSVIGNKTVEAMQSGIYVGFAGQVDGVVNRMAELADDDPDVTVIATGGLAPMVLGESS 240
QY 241 VIDHEPWLTLMLGLRLVYERNVSRM 265
241 VIDHEPWLTLMLGLRLVYERNVSRM 265

RESULT 2
AAU91152
ID AAU91152 standard; Protein; 272 AA.
XX AC AAU91152;
XX 05-JUN-2002 (first entry)
XX Mycobacterium tuberculosis pantothenate kinase Coax.
XX Pantothenate kinase; Coax; antibiotic; antimicrobial;
XX pantothenate kinase modulator; coenzyme A; bactericidal compound.
XX Mycobacterium tuberculosis.
XX WO200216601-A2.
XX 28-FEB-2002.
XX 24-AUG-2001; 2001WO-US26531.
XX 24-AUG-2000; 2000US-227860P.
XX 20-MAR-2001; 2001US-0813453.
XX (OMNI-) OMNIGENE BIOPRODUCTS INC.
XX Yocum RR, Patterson TA;
XX WPI; 2002-269358/31.
XX Identifying potential antibiotic or antimicrobial agent, comprises
XX contacting composition comprising pantothenate kinase (Coax) protein

```

```

PT with test compound and identifying inhibitor of the Coax protein
XX Claim 10; Page 70-71; 128pp; English.
XX The invention describes assays for identifying a (potential) antibiotic
XX CC comprising contacting an assay composition comprising a pantothenate
XX CC kinase (Coax) protein with a test compound, and determining the ability
XX CC of the test compound to inhibit the activity of the Coax protein, an
XX CC essential enzyme for the production of coenzyme A. Coax protein is a
XX CC valuable target for identifying bactericidal compounds. Coax modulating
XX CC agents can be used in an infectious animal model to determine the
XX CC efficacy, toxicity, or side effects of treatment with such an agent. This
XX CC is the amino acid sequence of a pantothenate kinase (Coax) protein
XX CC described in the invention.
XX SQ Sequence 272 AA;

Query Match 52.0%; Score 698.5; DB 23; Length 272;
Best Local Similarity 51.5%; Pred. No. 1.2e-59;
Matches 138; Conservative 53; Mismatches 64; Indels 13; Gaps 3;

QY 1 MLLTIDVGNTHVTLGLFDG----EDIVEHWRISTDSRRTADELAVLLOGLMGHPLLGDE 56
Db 1 MLLTIDVGNTHVTLGLFDG----EDIVEHWRISTDSRRTADELAVLLOGLMGHPLLGDE 56
QY 57 LGDGDGTAICATVPSVLHELREVTTRYGDPVAVLVEPGVKTGPILTDHPKVEGADRI 116
54 DSERLTGTAAALSTVPSVLHELREVTTRYGDPVAVLVEPGVKTGPILTDHPKVEGADRI 113
QY 117 INAAVAVELYGGPAIVDFGTATTDDAVSARGEYIGGVIAPEGIEISVEALGVKGAQLRK 176
114 VNCLAAVDFRKAALVDFGSSICVDVWSAKGEFLGGAIPGVQVSSDAARSAALRRV 173
QY 177 EVARPSVIGKNTVEAMQSGIYVGFAGQVDGVVNRMAELADDDPDVTVIATGGLAPM 234
174 ELARPSVIGKNTVEAMQSGIYVGFAGQVDGVVNRMAELADDDPDVTVIATGGLAPM 233
QY 235 VLGESSVIDEHEPWLTLMLGLRLVYERNV 262
234 LIPELHTVDHYDQHLTLQGLRLVFRNL 261

RESULT 3
AAG81225
ID AAG81225 standard; Protein; 272 AA.
XX AC AAG81225;
XX 04-SEP-2001 (first entry)
XX Mycobacterium tuberculosis potential drug target protein SEQ ID 276.
XX Drug target; growth; organism viability; characterisation.
XX Mycobacterium tuberculosis.
XX WO200135317-A1.
XX 17-MAY-2001.
XX 13-NOV-2000; 2000WO-US31152.
XX 12-NOV-1999; 99US-0165086.
XX 12-NOV-1999; 99US-0165124.
XX 01-FEB-2000; 2000US-0179531.
XX (REGC ) UNIV CALIFORNIA.
XX Eisenberg D, Rotstein SH, Marcotte EM;
XX WPI; 2001-329193/34.
XX DR N-PSDB; AAH52076.
XX

```


XX PA (OMNI-) OMNIGENE BIOPRODUCTS INC.
XX PI Yocum RR, Patterson TA;
XX DR WPI; 2002-269358/31.
XX DR N-PSDB; ABK54192.
XX PT Identifying potential antibiotic or antimicrobial agent, comprises
XX PT contacting composition comprising pantothenate kinase (Coax) protein
XX PT with test compound and identifying inhibitor of the Coax protein
XX PS Claim 10; Page 100; 128pp; English.
XX CC The invention describes assays for identifying a (potential) antibiotic
XX CC comprising contacting an assay composition comprising a pantothenate
XX CC kinase (Coax) protein with a test compound, and determining the ability
XX CC of the test compound to inhibit the activity of the Coax protein, an
XX CC essential enzyme for the production of coenzyme A. Coax protein is a
XX CC valuable target for identifying bactericidal compounds. Coax modulating
XX CC agents can be used in an infectious animal model to determine the
XX CC efficacy, toxicity, or side effects of treatment with such an agent. This
XX CC is the amino acid sequence of a pantothenate kinase (Coax) protein
XX CC described in the invention.
XX SQ Sequence 254 AA;
Query Match 49.3%; Score 662; DB 23; Length 254;
Best Local Similarity 49.8%; Pred. No. 3.8e-56;
Matches 132; Conservative 55; Mismatches 64; Indels 14; Gaps 4;
QY 1 MLLTIDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLLQGLMGHPLGLDELG-- 58
DB 1 MILVIDGNTNVLGVYQDETLVHWRILATSRQKTEDEYAMTVRSLF-----DHAGLIQ 53
QY 59 -DGIDGTAICATVPSVLHRELREYRYGDPVAVLVEPGVKTGVPILTDHPKEVCADRII 117
DB 54 FQDIDGIVISSVPPMFSQMKCKYHVTTP-MIIGIGITGLNLIKYNPREVGADRIV 112
QY 118 NAAVELYGGPAIVDFGTATFDVAVSARGEYIGGVIAAPGIEISVEALGVKAQLRKIE 177
DB 113 NAAVAIELYGPVAVDGTATTCLINEKQYAGGVIAAPGIMSTEALYHRASKLPRIE 172
QY 178 VARPSVIGKNTVAMQSGIVYGAGQVDGVNVRMARELADDDVTVIATGGLAPMVLG 237
DB 173 IAKPKQVVGNTIDSMQSGIFGVSVQDGVVVRKMAQAESEP---KVIATGGLAKLIGT 229
238 ESSVIDEHEPWLTLMLGLRLVYERNV 262
230 ESETIDVDSFTLTKGLQLIYKKNV 254
RESULT 6
AAU01243
ID AAU01243 standard; Protein; 258 AA.
AC AAU01243;
XX 18-JUL-2001 (first entry)
XX B. subtilis novel pantothenate kinase encoded by the gene coax.
XX Pantothenate kinase; coax; pantothenate biosynthesis; vitamin B5;
XX nutritional supplement; panto-compound; pantoate.
XX Bacillus subtilis.
XX WO200121772-A2.
XX 29-MAR-2001.
XX 21-SEP-2000; 2000WO-US25993.
XX

PR 21-SEP-1999; 99US-0400494.
PR 07-JUN-2000; 2000US-0210072.
PR 28-JUL-2000; 2000US-0221836.
PR 24-AUG-2000; 2000US-0227860.
XX (OMNI-) OMNIGENE BIOPRODUCTS.
XX Yocum RR, Patterson TA, Hermann T, Pero JG;
XX WPI; 2001-218644/22.
XX N-PSDB; AAS00984.
XX PT New recombinant microorganism which overexpress a Bacillus subtilis
XX PT pantothenate biosynthetic enzyme, useful for the high yield production
XX PT of panto-compounds such as pantothenate and pantoate
XX PS Example 14; Fig 23; 292pp; English.
XX CC The sequence represents a novel B. subtilis pantothenate kinase (encoded
XX CC by gene coax), an enzyme of the pantothenate biosynthetic pathway.
XX CC Pantothenate, also known as vitamin B5, is used as a nutritional
XX CC supplement in mammals and humans. The invention concerns methods of
XX CC producing recombinant microorganisms overexpressing at least one Bacillus
XX CC subtilis pantothenate biosynthetic enzyme. The microorganisms and methods
XX CC of producing them are useful for producing a panto-compound such as
XX CC pantothenate or pantoate, which is a nutritional requirement for
XX CC livestock and humans. The methods are also useful for the identification
XX CC of pantothenate kinase modulators. Panto-compounds are produced at a
XX CC significantly higher yield than prior art methods and can be produced
XX CC independent of the need to feed precursors which decreases expense.
XX SQ Sequence 258 AA;
Query Match 48.8%; Score 655; DB 22; Length 258;
Best Local Similarity 51.0%; Pred. No. 1.9e-55;
Matches 133; Conservative 46; Mismatches 74; Indels 8; Gaps 3;
QY 1 MLLTIDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLLQGLMGHPLGLDELG 60
DB 1 LLLVIDGNTNVLGVYHGDGKLEYHWRITSRKTEDEFGMLRSLFDHSGLMFEQ---- 56
QY 61 IDGTAICATVPSVLHRELREYRYGDPVAVLVEPGVKTGVPILTDHPKEVGADRINAV 120
DB 57 IDGIISSVPPIMFALERMCTKYPHIEPQI-VGPGMKTGLNLIKYNPREVGADRINAV 115
QY 121 AAVELYGGPAIVDFGTATFDVAVSARGEYIGGVIAAPGIEISVEALGVKAQLRKIEVAR 180
DB 116 AAHLIYGNPLIVDFGTATTTCYIDENKOYMGGAIPAGITISTEALYSRAAKLPRIETR 175
QY 181 PRSVIGKNTVAMQSGIVYGAGQVDGVNVRMARELADDDVTVIATGGLAPMVLGESS 240
DB 176 PDNIIGKNTVAMQSGILFYGQVGEIVKRMKWA---KQDLKVIATGGLAPLIANESD 232
QY 241 VIDEHEPWLTLMLGLRLVYERN 261
DB 233 CIDIVDPETLTKGLELIYERN 253
RESULT 7
AAU91149
ID AAU91149 standard; Protein; 258 AA.
XX AAU91149;
XX AAU91149;
XX 05-JUN-2002 (first entry)
XX Bacillus subtilis pantothenate kinase Coax.
XX Pantothenate kinase; Coax; antibiotic; antimicrobial;
XX pantothenate kinase modulator; coenzyme A; bactericidal compound.
XX Bacillus subtilis.
XX

PN WO200216601-A2.
 XX 28-FEB-2002.
 PD 24-AUG-2001; 2001WO-US26531.
 XX 24-AUG-2000; 2000US-227860P.
 PR 20-MAR-2001; 2001US-0813453.
 XX (OMNI-) OMNIGENE BIOPRODUCTS INC.
 XX Yocum RR, Patterson TA;
 XX WPI; 2002-269358/31.
 DR N-PSDB; ABK54168.
 XX Identifying potential antibiotic or antimicrobial agent, comprises
 PT contacting composition comprising pantothenate kinase (Coax) protein
 PT with test compound and identifying inhibitor of the Coax protein -
 XX Claim 10; Page 67-68; 128pp; English.
 XX The invention describes assays for identifying a (potential) antibiotic
 CC comprising contacting an assay composition comprising a pantothenate
 CC kinase (Coax) protein with a test compound, and determining the ability
 CC of the test compound to inhibit the activity of the Coax protein, an
 CC essential enzyme for the production of coenzyme A. Coax protein is a
 CC valuable target for identifying bactericidal compounds. Coax modulating
 CC agents can be used in an infectious animal model to determine the
 CC efficacy, toxicity, or side effects of treatment with such an agent. This
 CC is the amino acid sequence of a pantothenate kinase (Coax) protein
 CC described in the invention.
 XX Sequence 258 AA;
 SQ Query Match 48.8%; Score 655; DB 23; Length 258;
 Best Local Similarity 51.0%; Pred. No. 1.9e-55;
 Matches 133; Conservative 46; Mismatches 74; Indels 8; Gaps 3;
 QY 1 MLTIDVGNTHVGLFDGEDIVEHWRISTDSRRTADELAVLLQGLMGHPLGLDELGDG 60
 DB 1 LLLVIDVGNTHVGLVYHDKLEYHWRISTSRHKTDEFGMLRSLFDHSLMFEQ--- 56
 QY 61 IDGIAICATVPSVLHRELVRTRYGDPVAVLVEPGVKGVPIILTDHPKEVGADRIINAV 120
 DB 57 IDGIISSVPPIMFALERMCKYFIEPQI-VGPGMKTGLNKNYDNPKQVGDRIINAV 115
 QY 121 AAVELYGGPAIVVDFGTATTDAVSARGEYIGGVIAAPGIEISVEALGVGAOLRKEVAR 180
 DB 116 AAHLXGNPLVVDFTATTTCYIDENKQIMGAIAPIGTISTEALYSRAKLPRIETR 175
 QY 181 PRSVIGKNTVEAMQSGIVYGFAGQVGVNRMARELADDDPDVTVIATGGLAPMVLGESS 240
 DB 176 PDNIIGKNTVSAQSGILGYGVQVGVGVKRMKQA---KQDLKVIATGGLAPLIANESD 232
 QY 241 VIDEHEPWLTLMLGLRLVYERN 261
 DB 233 CIDIVDPFTLKLGLLEIYERN 253
 RESULT 8
 AAU91175
 ID AAU91175 standard; Protein; 256 AA.
 XX AC AAU91175;
 XX DT 05-JUN-2002 (first entry)
 XX DE Pantothenate kinase (Coax) #13.
 XX DE Pantothenate kinase; Coax; antibiotic; antimicrobial;
 KW Pantothenate kinase modulator; coenzyme A; bactericidal compound.
 XX

OS Clostridium difficile.
 XX WO200216601-A2.
 PN 28-FEB-2002.
 XX 24-AUG-2001; 2001WO-US26531.
 XX 24-AUG-2000; 2000US-227860P.
 PR 20-MAR-2001; 2001US-0813453.
 XX (OMNI-) OMNIGENE BIOPRODUCTS INC.
 XX Yocum RR, Patterson TA;
 PI WPI; 2002-269358/31.
 DR N-PSDB; ABK54196.
 XX Identifying potential antibiotic or antimicrobial agent, comprises
 PT contacting composition comprising pantothenate kinase (Coax) protein
 PT with test compound and identifying inhibitor of the Coax protein -
 XX Claim 6; Page 105; 128pp; English.
 XX The invention describes assays for identifying a (potential) antibiotic
 CC comprising contacting an assay composition comprising a pantothenate
 CC kinase (Coax) protein with a test compound, and determining the ability
 CC of the test compound to inhibit the activity of the Coax protein, an
 CC essential enzyme for the production of coenzyme A. Coax protein is a
 CC valuable target for identifying bactericidal compounds. Coax modulating
 CC agents can be used in an infectious animal model to determine the
 CC efficacy, toxicity, or side effects of treatment with such an agent. This
 CC is the amino acid sequence of a pantothenate kinase (Coax) protein
 CC described in the invention.
 XX Sequence 256 AA;
 SQ Query Match 47.9%; Score 643.5; DB 23; Length 256;
 Best Local Similarity 49.8%; Pred. No. 2.4e-54;
 Matches 130; Conservative 53; Mismatches 71; Indels 7; Gaps 4;
 QY 1 MLTIDVGNTHVGLFDGEDIVEHWRISTDSRRTADELAVLLQGLMGHPLGLDELGDG 60
 DB 1 MLVFDVGNTHVGLVYHDKLEYHWRISTSRHKTDEFGMLRSLFDHSLMFEQ--- 56
 QY 61 IDGIAICATVPSVLHRELVRTRYGDPVAVLVEPGVKGVPIILTDHPKEVGADRIINAV 120
 DB 57 IDDVIISSVPPIMFALERMCKYFIEPQI-LIVPGMKTGLNKNYDNPKQVGDRIINAV 115
 QY 121 AAVELYGGPAIVVDFGTATTDAVSARGEYIGGVIAAPGIEISVEALGVGAOLRKEVAR 180
 DB 116 AGIEKYGAPSLVDFGTATTTCYIDENKQIMGAIAPIGTISTEALYSRAKLPRIELAK 175
 QY 181 PRSVIGKNTVEAMQSGIVYGFAGQVGVNRMARELADDDPDVTVIATGGLAPMVLGESS 240
 DB 176 PGMTICKSTVSAQSGIIVYGVGLVDKIISIMKEL--NCDVVKVIATGGLAKLIASETK 233
 QY 241 VIDEHEPWLTLMLGLRLVYERN 261
 DB 234 SIDYVDGFLTLGLRIIYERN 254
 RESULT 9
 AAU91170
 ID AAU91170 standard; Protein; 262 AA.
 XX AC AAU91170;
 XX DT 05-JUN-2002 (first entry)
 XX DE Pantothenate kinase (Coax) #8.
 XX DE Pantothenate kinase; Coax; antibiotic; antimicrobial;
 KW

KW pantothenate kinase modulator; coenzyme A; bactericidal compound.
 XX Bacillus anthracis.
 OS WO200216601-A2.
 XX 28-FEB-2002.
 XX 24-AUG-2001; 2001WO-US26531.
 XX 24-AUG-2000; 2000US-227860P.
 XX 20-MAR-2001; 2001US-0813453.
 XX (OMNI-) OMNIGENE BIOPRODUCTS INC.
 XX Yocum RR, Patterson TA;
 XX WPI; 2002-269358/31.
 XX N-PSDB; ABK54191.
 XX
 XX Identifying potential antibiotic or antimicrobial agent, comprises
 XX contacting composition comprising pantothenate kinase (Coax) protein
 XX with test compound and identifying inhibitor of the Coax protein -
 XX
 XX Claim 8; Page 98-99; 128pp; English.
 XX
 XX The invention describes assays for identifying a (potential) antibiotic
 XX comprising contacting an assay composition comprising a pantothenate
 XX kinase (Coax) protein with a test compound, and determining the ability
 XX of the test compound to inhibit the activity of the Coax protein, an
 XX essential enzyme for the production of coenzyme A. Coax protein is a
 XX valuable target for identifying bactericidal compounds. Coax modulating
 XX agents can be used in an infectious animal model to determine the
 XX efficacy, toxicity, or side effects of treatment with such an agent. This
 XX is the amino acid sequence of a pantothenate kinase (Coax) protein
 XX described in the invention.
 XX
 XX Sequence 262 AA;
 XX
 XX Query Match 47.5%; Score 638; DB 23; Length 262;
 XX Best Local Similarity 47.9%; Pred. No. 8.6e-54;
 XX Matches 126; Conservative 55; Mismatches 74; Indels 8; Gaps 3;
 QY 1 MLLTIDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDLGDG 60
 Db 1 MIFVLDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDLGDG 60
 61 IDGTAICATVPSVLHELREVTTRYGYDPAVLVPEGVKTPILTDHPKEVGADRIINAV 120
 57 VKGIIVSSVPPIMFALERMCEKFKIKPLV-VGPGIKTGNIKYENPREVGADRIINAV 115
 121 AAELYGGPAIVDFGTATTFDAYSARGEYIGGVITAPGIEISVEALGVKGAOLRKIEVAR 180
 116 AGIHLGSLIIVDFGTATTTCYINEEKHYGGVITPGIMISAELYSRAKLPRIETK 175
 QY 181 PRSVIGKNTVEMASGIVYFAGQVGVNRMARELADDPDVTVIATGGLAPMYLGESS 240
 Db 176 PSSVVGKNTVMSAGSILYGVGVGVGIVKRMKEAKQEP---KYIATGGLAKLISESN 232
 QY 241 VIDHEPWLTLMLGLRLVYERNVS 263
 Db 233 VIDVDPFLTLKGLYMLYERN 255
 XX
 XX RESULT 10
 XX AAU91154
 XX ID AAU91154 standard; Protein; 255 AA.
 XX AC AAU91154;
 XX XX
 XX 05-JUN-2002 (first entry)
 XX DE Geobacter sulfurreducens pantothenate kinase Coax.

XX Pantothenate kinase; Coax; antibiotic; antimicrobial;
 KW pantothenate kinase modulator; coenzyme A; bactericidal compound.
 XX Geobacter sulfurreducens.
 OS WO200216601-A2.
 XX 28-FEB-2002.
 XX 24-AUG-2001; 2001WO-US26531.
 XX 24-AUG-2000; 2000US-227860P.
 XX 20-MAR-2001; 2001US-0813453.
 XX (OMNI-) OMNIGENE BIOPRODUCTS INC.
 XX Yocum RR, Patterson TA;
 XX WPI; 2002-269358/31.
 XX
 XX Identifying potential antibiotic or antimicrobial agent, comprises
 XX contacting composition comprising pantothenate kinase (Coax) protein
 XX with test compound and identifying inhibitor of the Coax protein -
 XX
 XX Claim 10; Page 72-73; 128pp; English.
 XX
 XX The invention describes assays for identifying a (potential) antibiotic
 XX comprising contacting an assay composition comprising a pantothenate
 XX kinase (Coax) protein with a test compound, and determining the ability
 XX of the test compound to inhibit the activity of the Coax protein, an
 XX essential enzyme for the production of coenzyme A. Coax protein is a
 XX valuable target for identifying bactericidal compounds. Coax modulating
 XX agents can be used in an infectious animal model to determine the
 XX efficacy, toxicity, or side effects of treatment with such an agent. This
 XX is the amino acid sequence of a pantothenate kinase (Coax) protein
 XX described in the invention.
 XX
 XX Sequence 255 AA;
 XX
 XX Query Match 46.8%; Score 629; DB 23; Length 255;
 XX Best Local Similarity 51.1%; Pred. No. 6.3e-53;
 XX Matches 134; Conservative 42; Mismatches 76; Indels 10; Gaps 5;
 QY 1 MLLTIDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDLG-D 59
 Db 1 MLLTIDVGNTHVTLGLFDGEDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDLG 55
 QY 60 GIDGTAICATVPSVLHELREVTTRYGYDPAVLVPEGVKTPILTDHPKEVGADRIINA 119
 Db 56 QIRAVIISVVPPLTGVLERLSLGYFGMRPLV-VGPGIKTGMPIQYDNPREVGADRIINA 114
 QY 120 VAAELYGGPAIVDFGTATTFDAYSARGEYIGGVITAPGIEISVEALGVKGAOLRKIEVA 179
 Db 115 VAGYEKTYSLIIVDFGTATTFDYNNRKEGCGGALPGLVISTEALFORASKLPVDII 174
 QY 180 PRSVIGKNTVEMASGIVYFAGQVGVNRMARELADDPDVTVIATGGLAPMYLGES 239
 Db 175 RPSAIARTVNSMOAGIYGVGLVDEIVTRMKAESKDAP---RVIATGGLASLIAPES 231
 QY 240 SVIDHEPWLTLMLGLRLVYERN 261
 Db 232 KTI EAEEYTLTLEGLRIILYERN 253
 XX
 XX RESULT 11
 XX ABB47661
 XX ID ABB47661 standard; Protein; 259 AA.
 XX AC ABB47661;
 XX XX
 XX 05-FEB-2002 (first entry)
 XX DE

DE Listeria monocytogenes protein #365.

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

KW vitamin B12; bacterial infection; disease.

XX Listeria monocytogenes.

XX WO200177335-A2.

XX 18-OCT-2001.

XX 11-APR-2001; 2001WO-FR01118.

XX 11-APR-2000; 2000FR-0004629.

XX (INSP) INST PASTEUR.

XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Deboux P;

PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;

PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;

PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;

PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;

PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;

PI Rose M, Voss H;

XX WPI; 2002-010914/01.

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment

PT and prevention of Listeria and related bacterial infections, and

PT related polypeptides

XX Claim 6; SEQ ID No 366; 192pp; French.

XX The present invention relates to the genome sequence of Listeria

CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of

CC it are useful for selecting probes and primers for detecting genes in L.

CC monocytogenes and related organisms, and for studying genetic

CC polymorphisms and other genomes. The present sequence is a protein

CC encoded by the genome sequence of the present invention. Proteins

CC expressed from the genome sequence are useful for raising specific

CC antibodies, identification of L. monocytogenes and related organisms, and

CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin

CC B12. The genome sequence and proteins encoded by it are also useful for

CC selecting compounds that regulate gene expression and cell replication

CC and modulate L. monocytogenes-related diseases. In addition, the genome

CC sequence and proteins encoded by it are useful in pharmaceutical and

CC vaccines compositions for the treatment or prevention of infections by L.

CC monocytogenes and related organisms.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 259 AA;

Query Match 44.4%; Score 596; DB 23; Length 259;

Best Local Similarity 44.8%; Pred. No. 1e-49;

Matches 117; Conservative 57; Mismatches 79; Indels 8; Gaps 3;

QY 1 MLLTIDVGNTHTVLGLFDGEDIVHWRISTDSRTADELAVLLQGLMGHPHLLGDELGDG 60

DB 1 MLIVDVGNTNCVGVYKOKRLKHWMTTDRHTSDLGNTLVNFFSYANLTPSD---- 56

QY 61 IDGIAICATVPSVHLEHRETVRYGDPVAVLVEPGVKTGVPILTDHPKEVGADRIINAV 120

DB 57 IQGIISSVPPIMHMETMVCVRIF-NIRPLVGPGLTKGLNLUKVDNPREIGSDRIINAV 115

QY 121 AAELYGGPAIVVDFGTATTFDAVSARGEYIGGVIAFGIETSVLEALGVKAQRLKIEVAR 180

DB 116 AASEEYGTPIVVDFGTATTFYCIDESGVYGGAIAPGIMISTEALYNRAKLPRVDIAE 175

QY 181 PRVIGKNTVEAMOSGIVYGFACGVGVNMRARELADDDPDDVTVIATGGLAPMVLGESS 240

DB 116 AAILHGNPLIVDFGTATTCYIDENKOYMGGAIPAGTITSTEALYSRAAKPLRIEITR 175

DB 176 SSQIIGKSTVSSMQAGIFYGVQCEGIIAEMKKQSNASP---VVVATGGLARMITEKSS 232

QY 241 VIDEHEPWLTLMLGLRLVYERN 261

DB 233 AVDILDPELTLKGLLELYRN 253

RESULT 12

AAU91163

ID AAU91163 standard; Protein; 233 AA.

XX AC AAU91163;

XX 05-JUN-2002 (first entry)

DE Pantothenate kinase (Coax) #1.

XX Pantothenate kinase; Coax; antibiotic; antimicrobial;

KW pantothenate kinase modulator; coenzyme A; bactericidal compound.

XX Bacillus subtilis.

OS WO200216601-A2.

PN 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US26531.

PF 24-AUG-2000; 2000US-227860P.

PR 20-MAR-2001; 2001US-0813453.

XX (OMNI-) OMNIGENE BIOPRODUCTS INC.

PA Yocum RR, Patterson TA;

XX WPI; 2002-269358/31.

DR N-PSDB; ABK54169.

XX Identifying potential antibiotic or antimicrobial agent, comprises

PT contacting composition comprising pantothenate kinase (Coax) protein

PT with test compound and identifying inhibitor of the Coax protein

XX Disclosure; Page 81-82; 128pp; English.

XX The invention describes assays for identifying a (potential) antibiotic

CC comprising contacting an assay composition comprising a pantothenate

CC kinase (Coax) protein with a test compound, and determining the ability

CC of the test compound to inhibit the activity of the Coax protein, an

CC essential enzyme for the production of coenzyme A. Coax protein is a

CC valuable target for identifying bactericidal compounds. Coax modulating

CC agents can be used in an infectious animal model to determine the

CC efficacy, toxicity, or side effects of treatment with such an agent. This

CC is the amino acid sequence of a pantothenate kinase (Coax) protein

CC described in the invention.

XX Sequence 233 AA;

Query Match 41.3%; Score 554.5; DB 23; Length 233;

Best Local Similarity 50.5%; Pred. No. 9.8e-46;

Matches 111; Conservative 39; Mismatches 65; Indels 5; Gaps 2;

QY 1 MLLTIDVGNTHTVLGLFDGEDIVHWRISTDSRTADELAVLLQGLMGHPHLLGDELGDG 60

DB 1 MLLVIDGNTVTLGVYHDKLEYHWRISTSHKTEDEFGMILRSIFDHSGLMFEQ---- 56

QY 61 IDGIAICATVPSVHLEHRETVRYGDPVAVLVEPGVKTGVPILTDHPKEVGADRIINAV 120

DB 57 IDGIISSVPPIMHMETMVCVRIF-NIRPLVGPGLTKGLNLUKVDNPREIGSDRIINAV 115

QY 121 AAELYGGPAIVVDFGTATTFDAVSARGEYIGGVIAFGIETSVLEALGVKAQRLKIEVAR 180

DB 116 AAILHGNPLIVDFGTATTCYIDENKOYMGGAIPAGTITSTEALYSRAAKPLRIEITR 175

Db 175 RAGNRIVGTDVFSANQSGVFWGYISLIEGLV---ARIKAEQEPMTVIATGGVASLFEG 231
QY 238 ESSVIDEHEPWTLMGLRLVYERN 261
Db 232 ATDSIDHFDSDLTIRGLLEIYRN 255

RESULT 15

AAU91150
ID AAU91150 standard; Protein; 250 AA.

XX AC

XX AC

XX DT

XX DT

XX DE

XX DE

XX KW

XX KW

XX S

XX PN

XX PN

XX PD

XX PD

XX PF

XX PR

XX PR

XX PR

XX PR

XX PA

XX PI

XX PI

XX DR

XX DR

XX PT

XX PT

XX PS

XX PS

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX SQ

XX SQ

05-JUN-2002 (first entry)

Clostridium acetobutylicum pantothenate kinase Coax.

Pantothenate kinase; Coax; antibiotic; antimicrobial;
pantothenate kinase modulator; coenzyme A; bactericidal compound.

Clostridium acetobutylicum.

WO200216601-A2.

28-FEB-2002.

24-AUG-2001; 2001WO-US26531.

24-AUG-2000; 2000US-227860P.

20-MAR-2001; 2001US-0813453.

(OMNI-) OMNIGENE BIOPRODUCTS INC.

Yocum RR, Patterson TA;

WPI; 2002-269358/31.

Identifying potential antibiotic or antimicrobial agent, comprises
contacting composition comprising pantothenate kinase (Coax) protein
with test compound and identifying inhibitor of the Coax protein.

Claim 10; Page 68-69; 128pp; English.

The invention describes assays for identifying a (potential) antibiotic
comprising contacting an assay composition comprising a pantothenate
kinase (Coax) protein with a test compound, and determining the ability
of the test compound to inhibit the activity of the Coax protein, an
essential enzyme for the production of coenzyme A. Coax protein is a
valuable target for identifying bactericidal compounds. Coax modulating
agents can be used in an infectious animal model to determine the
efficacy, toxicity, or side effects of treatment with such an agent. This
is the amino acid sequence of a pantothenate kinase (Coax) protein
described in the invention.

Sequence 250 AA;

Query Match 37.1%; Score 498.5; DB 23; Length 250;
Best Local Similarity 43.3%; Pred. No. 3e-40;
Matches 104; Conservative 50; Mismatches 73; Indels 13; Gaps 4;

QY 1 MLLTIDVGNTHVTLGFDGEDIVEHWRISTDSRRRADELAVLLQGLM---GMHPLLGDEL 57

Db 18 VILVDVGNNTNVLGTYNDTKLTAEWRLSTDVLSADEYGIQVMNLFQDQKLDPTL---- 73

QY 58 GDGIDGIAICAVPSVLHRELREVTTRYGDPVAVLVEPGVKTGVPILTDBPKREVGDRII 117

Db 74 ---VEGVIISSVVPNIMYKIRKFKINPLV-VGPGIKTGINKINPNKPEVGADRIY 129

QY 118 NAVAAYELVGGPAIVVDFGTATTFDAVSARGVYIGGVIAPGIEISVEALGVKGAQLRKIE 177

Db 130 NAVAAYELVGGPAIVVDFGTATTFDAVSARGVYIGGVIAPGIEISVEALGVKGAQLRKIE 189

QY 178 VAPRSVIGKNTVEAMQSGIVYGFAGQVDGVVNRMARLADDDPDVT--VIATGGLAPMV 235
Db 190 LIKPAYAICKNTISSIOSGIVRYLROVKYLFELKLENLPDGRRTTSLVLTATGGLAKLI 249

Search completed: June 24, 2003, 21:46:12
Job time : 28.207 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 21:36:04 ; Search time 9.66565 Seconds

(without alignments)
2635.685 Million cell updates/sec

Title: US-09-813-453A-4

Perfect score: 1343

Sequence: 1 MLLTIDVGNTHVLGLFDGE.....EPWLTMLGLRLYERNVSRM 265

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

ical number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1343	100.0	265	2 T36391	hypothetical prote
2	710.5	52.9	274	2 H86937	conserved hypotet
3	698.5	52.0	272	2 A70955	hypothetical prote
4	662	49.3	254	2 F83660	hypothetical prote
5	596	44.4	259	2 AF1102	conserved hypotet
6	596	44.4	273	2 E97293	probable transcrip
7	591	44.0	259	2 AF1464	conserved hypotet
8	554.5	41.3	233	2 S66100	conserved hypotet
9	517	38.5	261	2 B87489	transcription acti
10	409.5	30.5	262	2 E75516	conserved hypotet
11	342.5	25.5	246	2 D72320	conserved hypotet
12	291.5	21.7	273	2 D1326	conserved hypotet
13	181.5	13.5	262	2 F70165	conserved hypotet
14	180.5	13.4	592	2 H82031	probable biotin-la
15	177	13.2	592	2 B81009	Bira protein/Bvg a
16	174	13.0	257	2 S75559	hypothetical prote
17	167	12.4	276	2 A12292	hypothetical prote
18	139	10.3	242	2 A82637	conserved hypotet
19	138.5	10.3	248	2 H83111	hypothetical prote
20	134	10.0	267	2 I40327	baf protein - Bord
21	128.5	9.6	229	2 E70465	hypothetical prote
22	107.5	8.0	465	2 D83598	probable zinc prot
23	107	8.0	198	2 D95285	conserved hypotet
24	107	8.0	383	2 D86781	L-lactate oxidase
25	105.5	7.9	559	2 B75477	conserved hypotet
26	102.5	7.6	336	2 E72359	rod shape-determin
27	102.5	7.6	1036	2 JC5568	serine proteinase
28	102	7.6	781	2 T28750	hypothetical prote
29	101.5	7.6	293	2 AD0913	ribosomal protein

ALIGNMENTS

RESULT 1

T36391

hypothetical protein SCE94.31c - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jun-2000

C:Accession: T36391

R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.

submitted to the EMBL Data Library, April 1999

A:Reference number: 221573

A:Accession: T36391

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-265 <OLI>

A:Cross-references: EMBL:AL049628; PIDN:CAB40880.1; GSPDB:GN00070; SCODEB:SCE94.31c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCODEB:SCE94.31c

C:Superfamily: Streptomyces coelicolor hypothetical protein SCE94.31c

Query Match 100.0%; Score 1343; DB 2; Length 265;
Best Local Similarity 100.0%; Pred. No. 4.6e-92;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MLLTIDVGNTHVLGLFDGE	DIVEHWRISTDSRRTADELAVLLQGLMGHPLGLDGLG	60
Db	1	MLLTIDVGNTHVLGLFDGE	DIVEHWRISTDSRRTADELAVLLQGLMGHPLGLDGLG	60
Qy	61	IDGTAICATVPSVLHELREVT	RRYGGYPAVLVEPGVKGTPILTDHPKEVGADRIINAV	120
Db	61	IDGTAICATVPSVLHELREVT	RRYGGYPAVLVEPGVKGTPILTDHPKEVGADRIINAV	120
Qy	121	AAVELYGGPALVDFGTATT	FDVARSARGEYIGGVIAPIGIEISVEALGVKGAQLRKIEVAR	180
Db	121	AAVELYGGPALVDFGTATT	FDVARSARGEYIGGVIAPIGIEISVEALGVKGAQLRKIEVAR	180
Qy	181	PRSVIGKNTVEAMQSGIY	GFAGQVGVVNRMARLADPPDDVTVTATGGLAPMVLGESS	240
Db	181	PRSVIGKNTVEAMQSGIY	GFAGQVGVVNRMARLADPPDDVTVTATGGLAPMVLGESS	240
Qy	241	VIDEHEPWLTLMLGLRLY	ERNVSRM 265	
Db	241	VIDEHEPWLTLMLGLRLY	ERNVSRM 265	

RESULT 2

H86937

conserved hypothetical protein ML0232 [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C:Accession: H86937

R:Cooper, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Davies, R.N.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holt

riboflavin specifi
probable cleavage
aryldialkylphospha
rifamycin polyketi
nitrogenase (EC 1.
probable dihydroxy
probable integral
edta monooxygenase
nitrioltriacetate
pyruvate dehydrog
chitinase (EC 3.2.
chitinase (EC 3.2.
transposase - Rhiz
Mg(2+) transport A
embryonic protein

```

RESULT 5
AF1102
conserved hypothetical protein lmo0221 [imported] - Listeria monocytogenes (str.
C:Species: Listeria monocytogenes
C:date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AF1102
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.;
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournon,
ok, C.; Schlueder, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.;
A:title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MIMD:21537279; PMID:11679669

```

RESULT 8
S66100
conserved hypothetical protein yacB - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 18-Aug-2000
C:Accession: S66100; E69740
R:Ogasawara, N.; Nakai, S.; Yoshikawa, H.
DNA Res. 1, 1-14, 1994
A:Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis
A:Reference number: S65967; MUID:96051385; PMID:7584024
A:Accession: S66100
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-233 <OGA>
A:Cross-references: EMBL:D26185; NID:g467326; PIDN:RAA05305.1; PID:d1005847; PID:g468
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1994

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleith, S.; Schroter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, K.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-233 <KUN>
C:Cross-references: GB:299104; GB:AL009126; NID:92632267; PIDN:CAB11846.1; PID:ell182003;
C:Genetics:
A:Gene: yacB
A:Start codon: TTG
C:Superfamily: Streptomyces coelicolor hypothetical protein SCE94.31c

Query Match 41.3%; Score 554.5; DB 2; Length 233;
Best Local Similarity 50.5%; Pred. No. 7e-34;
Matches 111; Conservative 39; Mismatches 65; Indels 5; Gaps 2;

QY 1 MLTIDVGNTHVTLGLFDGEDIVHWRISTDSRTADELAVLLQGLMGHPLLGDELGDG 60
DB 1 MLLVLDVGNTHVTLGLFDGEDIVHWRISTDSRTADELAVLLQGLMGHPLLGDELGDG 56
QY 61 IDGTAICATVPSVLHRELVTRRYGDPVAVLVEPGVKTGPILTDHPKEVGADRINAV 120
DB 57 IDGIISSVPPINFALERCTKFEHPQI-VGPGMKTGUNKIYDNPKEVGADRINAV 115
QY 121 AAVELYGGPAIVDFGATTDFDAYSARGEYIGGVAPGIEISVEALGVKGAQLRKIEVAR 180
DB 116 AAHLVGNPLTVDFGATTDFDAYSARGEYIGGVAPGIEISVEALGVKGAQLRKIEVAR 175
QY 181 PRSVIGKNTVEAMQSGIVYGFAGQVDGVVNRARELADDDP 220
DB 176 PDNIIGNTVSAMQSGILFGYGVQGVGIVKRMKQAKQDP 215

SULT 9
7489
transcription activator, probable Baf family [Imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: B87489
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.G.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: B87489
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-261 <STO>
A:Cross-references: GB:AE005673; NID:g13423392; PIDN:AAK23910.1; GSPDB:GN00148
C:Genetics:
A:Gene: CCL935
C:Superfamily: Streptomyces coelicolor hypothetical protein SCE94.31c

Query Match 38.5%; Score 517; DB 2; Length 261;
Best Local Similarity 42.4%; Pred. No. 4.7e-31;
Matches 112; Conservative 50; Mismatches 90; Indels 12; Gaps 5;

QY 1 MLTIDVGNTHVTLGLFDGEDIVHWRISTDSRTADELAVLLQGLMGHPLLGDELGDG 59
DB 2 MLLAEQGNTHVTLGLFDGEDIVHWRISTDSRTADELAVLLQGLMGHPLLGDELGDG 56
QY 60 GIDGTAICATVPSVLHRELVTRRYGDPVAVLVEPGVKTGPILTDHPKEVGADRINA 119
DB 57 AIDAVIISVPPINFALERCTKFEHPQI-VGPGMKTGUNKIYDNPKEVGADRINAV 115
QY 120 AAVELYGGPAIVDFGATTDFDAYSARGEYIGGVAPGIEISVEALGVKGAQLRKIEVA 179
DB 116 IGAAMVYGPVTVDFGATTDFDAYSARGEYIGGVAPGIEISVEALGVKGAQLRKIEVA 175
QY 180 RP-RSVIGKNTVEAMQSGIVYGFAGQVDGVVNRARELADDDPDDVTVIATGGLAPMVLG 237
DB 176 RPAGNRIYGTDFVSAMQSGVFGYISLIEGLV---ARIKAERGEPTVIATGGLVFEFEG 232
QY 238 ESSVIDEHEPWLTLMLGLRLVYERN 261
DB 233 ATDSIDHEDSLTIRGLLEIYERN 256

RESULT 10
E75516
conserved hypothetical protein - Deinococcus radiodurans (strain RI)

C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Jun-2000
C:Accession: E75516

R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.C.
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans RI.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: E75516

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <WHI>

A:Cross-references: GB:AE001905; GB:AE000513; NID:96458144; PIDN:AAF10040.1; PID:964;
C:Genetics:
A:Gene: DR0461
A:Map position: 1
C:Superfamily: Streptomyces coelicolor hypothetical protein SCE94.31c

Query Match 30.5%; Score 409.5; DB 2; Length 262;
Best Local Similarity 37.9%; Pred. No. 4.1e-23;
Matches 97; Conservative 44; Mismatches 102; Indels 13; Gaps 4;

QY 2 LLTIDVGNTHVTLGLFDGEDIVHWRISTDSRTADELAVLLQGLMGHPLLGDELGDG 60
DB 6 LLAVDIGNTHVTLGLFDGEDIVHWRISTDSRTADELAVLLQGLMGHPLLGDELGDG 58

QY 61 IDGTAICATVPSVLHRELVTRRYGDPVAVLVEPGVKTGPILTDHPKEVGADRINAV 120
DB 59 IPRAVLSVAPPVGENVALKRFHMDAFVAVSENLPDVTVELDTPSGVADRINCLNF 118

QY 121 AAVELYGG--PAIVDFGATTDFDAYSARGEYIGGVAPGIEISVEALGVKGAQLRKIEV 178
DB 119 GAERYLGLDYAVVDFGTTNFVVGRRRRLGGLATGAQVSADALFARAALPRITL 178

QY 179 APRSVIGKNTVEAMQSGIVYGFAGQVDGVVNRARELADDDPDDVTVIATGGLAPMVLGE 238
DB 179 QAPETAIGKNTVHALQSLVFGYAEVMDGLLRIRAEI---PGEAVAVATGGSRTVOGI 235

QY 239 SSVIDEHEPWLTLMLGL 254
DB 236 QCEIDYDETLRLGL 251

RESULT 11
D72320
conserved hypothetical protein - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: D73320
R;Neelson, K.B.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickley
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: D73320
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-246 <ARN>
A;Cross-references: GB:AE001754; GB:AE000512; NID:g4981417; PIDN:AAD35964.1; PID:g498141
A;Experimental source: strain MS88
C;Genetics:
A;Gene: TM0883
C;Superfamily: Streptomyces coelicolor hypothetical protein SCE94.31c

Query Match 25.5%; Score 342.5; DB 2; Length 246;
Best Local Similarity 35.7%; Pred. No. 3.4e-18; Mismatches 98; Indels 23; Gaps 9;
Matches 92; Conservative 45;
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 20-Jun-2000
C;Accession: D71326
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 20-Jun-2000
C;Accession: D71326
R;Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: D71326
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-273 <COL>
A;Cross-references: GB:AE001220; GB:AE000520; NID:g3322705; PIDN:AAC65417.1; PID:g332271
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0431
C;Superfamily: Streptomyces coelicolor hypothetical protein SCE94.31c

Query Match 21.7%; Score 291.5; DB 2; Length 273;
Best Local Similarity 33.1%; Pred. No. 2.3e-14; Mismatches 108; Indels 25; Gaps 9;
Matches 89; Conservative 47;
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C;Accession: H82031
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C;Accession: H82031
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; P
Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajand
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 224;

QY 57 LGDGDGIDGATCATVPSVLHRELVTRRYGDPVAVLVEPGVKTGVPILTDHP--KEVGAD 114
Db 60 LRDAF----ISSVVPVLTTKTIADAVAQISG-VQPVFGPWAYEHLVRIPEVRAEIGD 114
QY 115 RIINAAVELYGGPAIVDFGTATFDVARSARGEYIGGVITAPGIEISVEALGVKGAQLR 174
Db 115 LVANAAVYHFRSACVVDCGTALTFTAVDCTGLIOGVATAPGLRTAVQSILHTTAQLP 174
QY 175 KIEVARPSRVIGKNTVEAMQSGIYVGFAGQVDGVVNMARELADDDPDVTVIATGGLAPM 234
Db 175 LVPLALPDSVLGKDTTHAVAGVVGRTFLVIRAMIAQCQKEIG---CRCAAVITGLSRL 231
QY 235 VLGESSVID--EHEPWTLMGL-----RLV 257
Db 232 F---SSEVDFPPIDAQLTSLGLAHRLV 257

RESULT 13

F70165
Conserved hypothetical protein BB0527 - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C;Accession: F70165
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; V
son, D.; Peterson, J.; Ravavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, E
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943; PMID:9403685
A;Accession: F70165
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-262 <KLE>
A;Cross-references: GB:AE001154; GB:AE000783; NID:g2688431; PIDN:AAC66882.1; PID:g26
A;Experimental source: strain B31

Query Match 13.5%; Score 181.5; DB 2; Length 262;
Best Local Similarity 23.6%; Pred. No. 2.9e-06; Mismatches 114; Indels 35; Gaps 7;
Matches 63; Conservative 55;
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 20-Jun-2000
C;Accession: D71326
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 20-Jun-2000
C;Accession: D71326
R;Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: D71326
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-273 <COL>
A;Cross-references: GB:AE001220; GB:AE000520; NID:g3322705; PIDN:AAC65417.1; PID:g332271
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0431
C;Superfamily: Streptomyces coelicolor hypothetical protein SCE94.31c

RESULT 14

H82031
probable biotin-[acetyl-CoA-carboxylase] ligase (EC 6.3.4.15) - Neisseria meningitidis
N;Contains: biotin-[acetyl-CoA-carboxylase] ligase (EC 6.3.4.15)
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C;Accession: H82031
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; P
Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajand
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 224;

A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: H82031
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-592 <PAR>
A:Cross-references: GB:AL12752; GB:AL157959; NID:97378778; PIDN:CA83659.1; PID:9737911
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: birA; NMA0357
C:Keywords: ligase

Query Match 13.4%; Score 180.5; DB 2; Length 592;

Best Local Similarity 27.0%; Pred. No. 9.1e-06;
Matches 67; Conservative 35; Mismatches 103; Indels 43; Gaps 8;

```
QY 28 ISTDSRTADELAVLLOQ-----LMGMHPL-----LGDELGDGID----- 62
DB 329 VSPKRDSEFLLLDGNSRLKNAWVNGTFAIVGSAPYRDLSPLGAEWAEEKVDGNVRI 388
DB 63 -GIAICATVPSVLHELREVTTRYGDVPALVPEGVKTVGPILTDHPKEVGADRIINAVA 121
DB 389 VGCAVCGEFKKA--QVQELARKLEWLPSSAQALGIRNHY----RHPEHSGDRWFNALG 442
QY 122 AVELYGGPAIVDFGTATFDVARGEYIGGVITAPGIEISVEALGVKGAQIRKIEVARP 181
DB 443 SRFRNACVVVSCGTAVTVDALDGDGHLGGTIMGPHLMKESLAVETANLNR--HAGK 500
QY 182 RSVIGKNTVEAMQSGIYVGFAGQVDGVVNRMARLADDPDDVTVIATGG-----LAP 233
DB 501 RYPPPTTGNVAVSGMDAVCGSVMMHGRL-KKTGAGKPDVVIITGGGAQVAEALPP 559
QY 234 VVLGESSV 241
DB 560 AFLAENTV 567
```

RESULT 15

B81009
BirA protein/Bvg accessory factor NMB2075 [Imported] - Neisseria meningitidis (strain MC58)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: B81009
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Vignani, A.; Tettelin, H.; et al.
Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
Reference number: A81000; MUID:20175755; PMID:10710307

Accession: B81009
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-592 <TET>
A:Cross-references: GB:AE002557; GB:AE002098; NID:97227332; PIDN:AAF42394.1; PID:9722733
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB2075

Query Match 13.2%; Score 177; DB 2; Length 592;

Best Local Similarity 30.5%; Pred. No. 1.7e-05;
Matches 60; Conservative 28; Mismatches 89; Indels 20; Gaps 6;

```
QY 56 ELGDG----IDGIALCATVPSVLHELREVTTRYGDVPALVPEGVKTVGPILTDHPKEVG 112
DB 380 EKADGNVRIVGCACVGEFKA--QVQELARKLEWLPSSAQALGIRNHY----RHPEHG 433
QY 113 ADRIINAAVELYGGPAIVDFGTATFDVARGEYIGGVITAPGIEISVEALGVKGAQ 172
DB 434 SDRWFNALGSRFRNACVVVSCGTAVTVDALDGDGHLGGTIMGPHLMKESLAVRTAN 493
QY 173 LRKIEVAPRSVIGKNTVEAMQSGIYVGFAGQVDGVVNRMARLADDPDDVTVIATGG-- 230
DB 494 LNR--HAGKRYPPPTTGNVAVSGMDAVCGSVMMHGRL-KKTGAGKPDVVIITGGGA 550
```

QY 231 -----LAPMVLGESSV 241
DB 551 AKVAEALPPAFLAENTV 567

Search completed: June 24, 2003, 22:02:59
Job time : 11.6657 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 21:33:11 ; Search time 4.60909 Seconds
(without alignments)
2384.688 Million cell updates/sec

Title: US-09-813-453A-4
Perfect score: 1343
Sequence: 1 MLLTIDVGNTHVTLGLDFGE.....EPWLTMLGLYVERNVRSM 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554.5	41.3	233	1 YACB_BACSU	P37564 bacillus su
2	134	10.0	267	1 BAF_BORPE	Q45338 bordetella
3	99.5	7.4	461	1 NIFN_RHOCA	P19077 rhodobacter
4	99.5	7.4	557	1 ILVD_AERPE	Q39768 aeropyrum p
5	99.5	7.4	705	1 MMLC_STRCO	O88022 streptomyce
6	97.5	7.3	472	1 KPYK_MYCIT	P94939 mycobacteri
7	97	7.2	820	1 CHIA_ALTSO	P32823 alteromonas
8	96.5	7.2	811	1 MUTS_THEAQ	O56215 thermus aqu
9	95.5	7.1	438	1 TUDA_AGRVI	Q44472 agrobacteri
10	95.5	7.1	472	1 KPYK_MYCTU	O06134 mycobacteri
11	95	7.1	279	1 TRPA_RHILO	Q98cn6 thizobium l
12	95	7.1	321	1 GLK_BACSU	P54495 bacillus su
13	93.5	7.0	242	1 HIS4_HALN1	Q9hn14 halobacteri
14	93.5	7.0	293	1 PRMA_ECOLI	P28637 escherichia
15	93	6.9	355	1 BUKL_CIOAB	Q45829 clostridium
16	93	6.9	639	1 YGFT_ECOLI	Q46820 escherichia
17	93	6.9	697	1 YN26_MYCTU	P71886 mycobacteri
18	92.5	6.9	525	1 GUAA_PSEAE	O9hxm5 pseudomonas
19	92	6.9	261	1 PYRH_MYCTU	Q10791 mycobacteri
20	91.5	6.8	1113	1 CARB_CORGL	P58939 corynebacte
21	91	6.8	1101	1 GUNC_CELFI	P14090 cellulomona
22	90.5	6.7	973	1 UVRA_RHIME	P56899 rhizobium m
23	90	6.7	284	1 LPXD_METCA	Q9aip8 methylococc
24	90	6.7	397	1 TYRB_SALTY	P74861 salmonella
25	90	6.7	937	1 SYL_METHH	O27552 methanobact
26	89.5	6.7	269	1 TRCL_STRCO	O68814 streptomyce
27	89.5	6.7	553	1 SYE_METHH	O26157 methanobact
28	89.5	6.7	708	1 TRP_NEUCR	P13228 neurospora
29	88.5	6.6	277	1 PANB_BACSU	P52966 bacillus su
30	88.5	6.6	449	1 TRB2_AERPE	Q9y9h2 aeropyrum p
31	88	6.6	379	1 AMPC_MORMO	P94958 morganella
32	87	6.5	517	1 MURE_MYCLE	O69556 mycobacteri
33	87	6.5	821	1 CLPC_PORPY	P51332 porphyra pu

ALIGNMENTS

RESULT 1

YACB_BACSU
ID YACB_BACSU STANDARD; PRT; 233 AA.
AC P37564; 1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yacB.
GN YACB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
Presecan E., Pujic P., Purnelle B., Rappoport G., Rey M., Reynolds S.,
Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadafé Y.,
Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
Sakaguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,
Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis.";
RL Nature 390:249-256(1997).
CC -!- SIMILARITY: SOME, TO B.PERTUSSIS BVG ACCESSORY FACTOR..

34 86.5 6.4 335 1 GPDA_ARCFU
35 86.5 6.4 429 1 YN05_MYCTU
36 86 6.4 269 1 THIG_CAUCR
37 86 6.4 327 1 PXA2_SALTI
38 85.5 6.4 265 1 Y309_HALN1
39 85.5 6.4 304 1 PHEA_AMEYE
40 85.5 6.4 349 1 RUVB_MYCLE
41 85.5 6.4 475 1 KPYK_CORGL
42 85.5 6.4 1018 1 HMW1_MYCPN
43 85 6.3 327 1 PXA2_SALTY
44 85 6.3 379 1 FTS2_MYCTU
45 85 6.3 427 1 SYH_MYCLE

O29390 archaeoglob
O50660 mycobacteri
Q9a746 caulobacter
P58716 salmonella
Q9hsb8 halobacteri
Q44104 amycolatops
P40833 mycobacteri
O46078 corynebacte
Q50365 mycoplasma
P58718 salmonella
O08378 mycobacteri
P46696 mycobacteri

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D26185; BAA05305.1; -
 CC EMBL; Z99104; CAB11846.1; -
 CC Subtilisin; BG10133; YacB.
 CC InterPro; IPR004619; Baf.
 CC Pfam; PF03309; Bvq_acc_factor; 1.
 CC TIGRfams; TIGR00671; baf; 1.
 CC Hypothetical protein; complete proteome.
 CC SEQUENCE 233 AA; 26217 MW; AAE96E732C15DF44 CRC64;

Query Match 41.3%; Score 554.5; DB 1; Length 233;
 Best Local Similarity 50.5%; Pred. No. 1.1e-33;
 Matches 111; Conservative 39; Mismatches 65; Indels 5; Gaps 2;
 QY 1 MLLTIDVGNTHVTLGLDGEDIVHWRISTDSRTADELAVLQGLMGHPHLLGDLG 60
 Db 1 MLLVIDVGNTHVTLGLVHDGKLEVHWRIETSRHKTDEFGMLRSLFDHSGLMFEQ 56
 QY 61 IDGTAICATVPSVLHREVTTRYGDVPAVLVPEGVTVGPIILTDHPKEVGADRIINAV 120
 Db 57 IDGIIISVVPPIPFALERMCTKTFHIEPQI-VGPGMKTGLNIRKYDNPKEVGADRIINAV 115
 QY 121 AAVELYGGPATVVDGFTATPDAYSARGEYIGGVIAPIGIELSVLEALGVKGQALRKIEVAR 180
 Db 116 AAHLXGNPLVDFGATTVCYIDENKQYMGGAIPAGITITSTALYSRAAKLPRIETR 175
 QY 181 PRSVIGKNTVEAMSGVIGYFAGQVGVGVNRMARELADDP 220
 Db 176 PDNIIGKNTVSAMSGILFVGVGVEGIVKRMKQAKQDP 215

RESULT 2

ID_BAF_BORPE STANDARD; PRT; 267 AA.
 AC Q45338; Q45373;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Bvq accessory factor.
 GN BAF.
 CC Bordetella pertussis.
 CC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 CC Bordetella.
 CC NCBI_TaxID=520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BP504;
 RX MEDLINE=95325323; PubMed=7601846;
 RA Deshazer D., Wood G.E., Friedman R.L.;
 RT "Identification of a Bordetella pertussis regulatory factor required
 RT for transcription of the pertussis toxin operon in Escherichia
 RT coli.";
 RL J. Bacteriol. 177:3801-3807 (1995).
 RN [2]
 RP SEQUENCE OF 1-38 FROM N.A.
 RC STRAIN-BP504;
 RX MEDLINE=96419162; PubMed=8821935;
 RA Allen A.G., Maskell D.J.;
 RT "The identification, cloning and mutagenesis of a genetic locus

RT required for lipopolysaccharide biosynthesis in Bordetella
 RT pertussis.";
 RL Mol. Microbiol. 19:37-52 (1996).
 CC -!- FUNCTION: ACTIVATES TRANSCRIPTION OF THE PERTUSSIS TOXIN OPERON IN
 CC A BVGAS-DEPENDENT MANNER. MAY INTERACT WITH THE ALPHA SUBUNIT OF
 CC RNA POLYMERASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U12020; AAA75361.1; -
 CC EMBL; AF016461; AAC68834.1; -
 CC EMBL; X90711; CAA62242.1; -
 CC InterPro; IPR004619; Baf.
 CC Pfam; PF03309; Bvq_acc_factor; 1.
 CC Transcription regulation; Activator.
 CC SEQUENCE 267 AA; 27845 MW; 336A615F67B57901 CRC64;
 Query Match 10.0%; Score 134; DB 1; Length 267;
 Best Local Similarity 25.0%; Pred. No. 0.0049;
 Matches 58; Conservative 33; Mismatches 79; Indels 62; Gaps 9;
 QY 1 MLLTIDVGNTHVTLGLDGEDIVHWRISTDSRTADELAVL 42
 Db 1 MIILIDSGNSRLKVGWDFDPAQAREPAPVAFDNLDLALGRW-LATLPRRQALGVN 59
 QY 43 LOGLMGHPHLLGDLGDLGDTGTAICATVPSVLHREVTTRYGDVPAVLVPE---GVKT 99
 Db 60 VAGL-----ARGEAIATL-----RAGCCDIRLWRAQPLAMGLRN 94
 QY 100 GVPILTDPHPKEVGADR---IINAAVAELYGGPAIVVDFGTATFADVASRGEVIGVIA 156
 Db 95 GY----RNPDLQGLADRWACWGVGLARQPSVHPHLLVASGTATLTIGDPNVPFGGLIL 150
 QY 157 PGIEISVEALGVKGQALRKIEVARPSVIGK---NTVEAMSGVIGYFAGQV 205
 Db 151 PGPMAMRGALAYGTAHLPLAD-----GLVADYPIDTHQAIASGIAAAQAAGAI 197

RESULT 3

ID_NIFN_RHOCA STANDARD; PRT; 461 AA.
 AC P19077;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nitrogenase iron-molybdenum cofactor biosynthesis protein nifN.
 GN NIFN.
 OS Rhodospirillum rubrum (Rhodospirillum rubrum).
 CC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum group;
 CC Rhodospirillum.
 CC NCBI_TaxID=1061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89313674; PubMed=2747620;
 RA Moreno-Vivian C., Schmeil M., Masepohl B., Arnold W., Klipp W.;
 RT "DNA sequence and genetic analysis of the Rhodospirillum rubrum
 RT nifN gene region: homology between nifX and nifB suggests
 RT involvement of NifX in processing of the iron-molybdenum cofactor.";
 RL Mol. Gen. Genet. 216:353-363 (1989).
 CC -!- FUNCTION: THIS PROTEIN MAY PLAY A ROLE IN THE BIOSYNTHESIS OF
 CC THE PROSTHETIC GROUP OF NITROGENASE (FEMO COFACTOR).
 CC -!- PATHWAY: Fe-Mo cofactor biosynthesis.
 CC -!- SIMILARITY: BELONGS TO THE NIFD/NIFK/NIFE/NIFN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch)

CC EMBL: X17433; CAA35473.1; -
 CC PIR: JE0030; JE0030.
 CC HSSP: P11347; LMO.
 CC InterPro: IPR000318; Nitrogenase_comp1.
 CC InterPro: IPR000510; Oxidized_nitrogen.
 CC Pfam: PF00148; oxidized_nitro; 1.
 CC TIGRfams: TIGR01285; nifN; 1.
 CC PROSITE: PS00699; NITROGENASE_1.1; 1.
 CC PROSITE: PS00090; NITROGENASE_1.2; FALSE_NEG.
 CC Nitrogen fixation.
 KW METAL 44
 FT METAL 44 FEMO COFACTOR LIGATION (POTENTIAL).
 SQ SEQUENCE 461 AA; 49450 MW; 13A1F3F3B7D52872 CRC64;

Query Match 7.4%; Score 99.5; DB 1; Length 461;
 Best Local Similarity 23.8%; Pred. No. 2.9;
 Matches 65; Conservative 36; Mismatches 85; Indels 87; Gaps 15;
 QY 22 IVEHWR--ISTDSRRTADEL--AVLLQGLMGHPLGLDGLDGTGTAICATVPSVLHEL 77
 DB 3 VLTHSRRLSTNPLKTSAPLGRAMAYLGIEGAVPLF-----HGAQG-----CTAFGVVHLV 53
 QY 78 RYRTRYGVDPVAVLPEGVKTVGPILTDPKEV-----GADRIINAAVAEVLGGPAIV 132
 DB 54 RH-----FKEAVPLQTTAMNEVSTLGGGEQIEAIDNIRKRANPKFI 96
 QY 133 VDEGTATTDAVSARGEYIGGVTAPEIETSVKQALRIEVARPSVTKNTVEA 192
 DB 97 GIASTALT-----ETRGEDIAG-----ELRAMQVRR-KDWGTVAVHV 133
 QY 193 M-----QSGIVYGAGVDGVNRMARELAD-DPD-----DVTVIAT 228
 DB 134 IYDFEGGQDGHAKAVEALVAVPTAERDPLRQVTLVPSCTTAEIDAVRMIRA 193
 QY 229 GGLAPVLGE-SSVIDEH--EPWL--TLMLGLRL 256
 DB 194 FGLSPVLPLDSTSLDGLSDWSHSLGGTRL 226

RESULT 4
 ILVD_AERPE STANDARD; PRT; 557 AA.
 AC QY988;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Dihydroxy-acid dehydratase (EC 4.2.1.9) (DAD).
 GN ILVD OR APE0013.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 OC Desulfurococcales; Aeropyrum.
 OX NCBI_Taxid=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kwarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 CC -1- CATALYTIC ACTIVITY: 2,3-dihydroxy-3-methylbutanoate = 3-methyl-2-oxobutanoate + H(2)O.
 CC -1- COFACTOR: BINDS 1 4FE-4S CLUSTER (POTENTIAL).
 CC -1- PATHWAY: Valine and isoleucine biosynthesis; fourth step.

CC -1- SIMILARITY: BELONGS TO THE ILVD / EDD FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch)

CC EMBL: AP000058; BAA78922.1; ALT_INIT.
 CC InterPro: IPR000581; ILVD_EDD_family.
 CC InterPro: IPR004404; ILVD.
 CC Pfam: PF00920; ILVD_EDD; 1.
 CC PRODOM: PD002691; ILVD_EDD_family; 1.
 CC TIGRfams: TIGR00110; ilvd; 1.
 CC PROSITE: PS00886; ILVD_EDD_1; 1.
 CC PROSITE: PS00887; ILVD_EDD_2; 1.
 CC Branched-chain amino acid biosynthesis; Lyase; Iron-sulfur; 4Fe-4S;
 KW Complete proteome. 126 IRON-SULFUR (4FE-4S) (POTENTIAL).
 FT METAL 126 IRON-SULFUR (4FE-4S) (POTENTIAL).
 FT METAL 198 IRON-SULFUR (4FE-4S) (POTENTIAL).
 SQ SEQUENCE 557 AA; 58973 MW; 05C16E34F742F0BE CRC64;

Query Match 7.4%; Score 99.5; DB 1; Length 557;
 Best Local Similarity 24.2%; Pred. No. 3.6;
 Matches 65; Conservative 33; Mismatches 70; Indels 101; Gaps 14;
 QY 88 VPAVLVPPGVKTPVILTDH--PREVGADRIINAVA----- 121
 DB 78 VPIVVDNGINMGTPGMYSLISRELIADTIEAQVASHGDFGVGCGDKTQPGIMMAM 137
 QY 122 -----AVELYGGPAI--VVDFTATTFDVSARGEYIGVVI-----APG- 158
 DB 138 ARLDLPISYIYGGTAEHGVLDGETVTVQSAPAVGAVLKLIDERLYEIEKAAMPTGCT 197
 QY 159 -----TEISVEALGVK-----GAQLRKIEVAR-----PRSVIG 186
 DB 198 CQGLFTANTWAILAEALGLSPLGSASPATSSERARELARAGALAVGLVETGLTPRRIL- 256
 QY 187 KNTVEAMOSGIV--YGFAGVDGVVN--RMARELA-----DDPD-----VTVIANGGLAP 233
 DB 257 --TYEAFYNAIVTLMAISGSTNAVHLLAIAREAGVKLALDDFDEASKRPVIA--ALAP 312
 QY 234 MVLGESSVIDEHEPWLTMGLRLYERNV 262
 DB 313 --AGKYMVDLHNVGAPVILRKLDRGL 339

RESULT 5
 MMLC_STRCO STANDARD; PRT; 705 AA.
 ID MMLC_STRCO
 AC O88022;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative membrane protein SC06666.
 GN SC06666 OR SC5A7.16C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_Taxid=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,

RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RL coelicolor A3(2).";
 RC Nature 417:141-147(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE MMPL FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL031107; CA919945.1;
 DR InterPro; IPR000731; HMGCR/patch_5TM.
 DR InterPro; IPR004869; MMPL.
 DR Pfam; PF03176; MMPL; 2.
 DR PROSITE; PS0156; SSD; 1.
 FT Hypothetical protein; transmembrane; Complete proteome.
 FT TRANSMEM 16 36 POTENTIAL.
 FT TRANSMEM 144 164 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 201 221 POTENTIAL.
 FT TRANSMEM 232 252 POTENTIAL.
 FT TRANSMEM 280 300 POTENTIAL.
 FT TRANSMEM 306 326 POTENTIAL.
 FT TRANSMEM 360 380 POTENTIAL.
 FT TRANSMEM 504 524 POTENTIAL.
 FT TRANSMEM 528 548 POTENTIAL.
 FT TRANSMEM 561 581 POTENTIAL.
 FT TRANSMEM 615 635 POTENTIAL.
 FT TRANSMEM 636 656 POTENTIAL.
 SQ SEQUENCE 705 AA; 73118 MW; B1935ACE0F75390E CRC64;

Query Match 7.4%; Score 99.5; DB 1; Length 705;
 Best Local Similarity 26.1%; Pred. No. 4.6;
 Matches 65; Conservative 26; Mismatches 89; Indels 69; Gaps 13;
 QY 30 TDSRTADELAVALQGMHPLLGDELGDGIDGIAICATV--PSVLHELRL----- 78
 DB 48 TESQVAEEH---LDGL-----DPAGDSITGVVEAAVADPAVRAEVRVADLREV 95
 QY 79 ----EVTRRYGDPVAVLVEPGVKGVPI-----LTDHPKEV-----GADRII-----NAVA 121
 DB 96 AGVAEVPDPY--ATPGTVAEDGRALVSVTLLEGGLDDDAEAAVDDAADRLHIGIDGSAVS 153
 DB 122 AVELYGGPAIVDFGTATTFDAVSARGEYI-----GGVIAPGIEISVEALGV 168
 DB 154 GVHVSGLLQOQLGERAQEDVKNA--ELISLPVVLVLLVVFGGLRAAGLPLLVAVAGI 211
 QY 169 KGAOLRKTIEARPSVIGKNTVEAMOSGIVYGFAGQVGVNVRMAR---ELADDPDDV-- 223
 DB 212 AGAFALAFGFSHVTDI-----SVYAIQVTTMLGLGLAVDIALMLVFRERHRHPDVVEA 267
 QY 224 ---TVIATG 229
 DB 268 VHVTVAAAG 276

RESULT 6
 KPYK_MYCIT
 ID KPYK_MYCIT STANDARD; PRT; 472 AA.
 AC P94939;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pyruvate kinase (EC 2.7.1.40) (PK).
 GN PYK OR PYKF.
 OS Mycobacterium intracellulare.
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1767;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Batty;
 RA Alavi M.R., Rouse D.A., Morris S.L.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate -> ADP + phosphoenolpyruvate.
 CC -1- COFACTOR: REQUIRES MAGNESIUM AND POTASSIUM.
 CC -1- PATHWAY: Glycolysis; final step.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PYRUVATE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U65430; AAB39214.1;
 DR HSSP; P14178; LEOT.
 DR InterPro; IPR001697; Pyruvate_Kinase.
 DR Pfam; PF0224; PK; 1.
 DR Pfam; PF02887; PK_C; 1.
 DR PRINTS; PRO1050; PYRUVTKNASE.
 DR ProDom; PD001009; Pyruvate_Kinase; 1.
 DR TIGRFAMs; TIGR01064; pyruv_Kin; 1.
 DR PROSITE; PS00110; PYRUVATE_KINASE; 1.
 KW Pyruvate; Transferase; Kinase; Glycolysis; Magnesium.
 FT ACT_SITE 218 218 BY SIMILARITY.
 FT METAL 220 220 MAGNESIUM (BY SIMILARITY).
 FT METAL 241 241 MAGNESIUM (BY SIMILARITY).
 FT METAL 242 242 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 472 AA; 50461 MW; A9D58F3E64E117D1 CRC64;
 Query Match 7.3%; Score 97.5; DB 1; Length 472;
 Best Local Similarity 22.3%; Pred. No. 4.2;
 Matches 62; Conservative 36; Mismatches 77; Indels 103; Gaps 14;
 QY 18 DGEI-----VEHWRISTDSRTADELAVALQGMHPLLGDE---LGDGIDGIA 65
 DB 87 DGETVRITVADCEGSHDRVSTYKKLAEDAAV-----GDRVLVDGKVLVDGLEG-- 138
 QY 66 ICATVPSVLHELRETRRYGYDVPVAVLVE-----DDVICTVVEGGPVSNKGISLPGMNVSA
 DB 139 -----DDVICTVVEGGPVSNKGISLPGMNVSA
 QY 111 ----VGADRI-----INAAVAVEL-----YGGPAIVDFGT-ATTFDAV 144
 DB 179 FALDGLGDLVALSFVRSPADVELVHEVMDRVGRVVRVIAKLEKPEAVDNLETIVLAFDAI 238
 QY 145 S-ARGEYIGGVITAPGIEISVEALGVKGAOLRKTIEARPSVIGKNTVEAMOSGIVYGFAG 203
 DB 239 MVARGDL-----GVLEPLEEVL--VOKRAIQMARENA---KPVIVATQ----- 277
 QY 204 QVDGVNVRMARLADDPDDTVTVIATGGLAPMVLGESSV 241
 DB 278 MLDMSIENSRRPTREASDVANAVLDGADAVMLSGTSSV 315
 RESULT 7
 CHIA_ALTSO
 ID CHIA_ALTSO STANDARD; PRT; 820 AA.
 AC P32823;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chitinase A precursor (EC 3.2.1.14) (CHI-A).
 GN CHIA.
 OS Alteromonas sp. (strain O-7).
 CC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;


```

QY 92 LVEPGVKTGVPILTDHPKVEGADRILNA-VAARELYGGPAIVVDFGTATTFDAVSARGEY 150
Db 703 -----LGLPRKLNLFVAREEAGGLVFY-----HQV 728
QY 151 IGGVIAPGIEISVEALGVKGAOLRIEVARPSRVIGKNTVEAMQSGIYVGFAGQVGVVN 210
Db 729 LPGPASKSYGVEVARM-----AGLPKEVVARALL-----QAMAA-----RRGALDAVLE 775
QY 211 RWARELADDDPDVT 224
Db 776 RL---LALDDPRLT 786

RESULT 9
TUD4_AGRVI STANDARD; PRT; 438 AA.
ID TUD4_AGRVI
AC Q44472;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative hydroxypyruvate reductase (EC 1.1.1.81).
GN TTUD.
OS Agrobacterium vitis (Rhizobium vitis).
OC Plasmid pTRAB4.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=373;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB4;
RX MEDLINE=96062236; PubMed=7592429;
RA Crouzet P., Otten L.;
RT "Sequence and mutational analysis of a tartrate utilization operon
from Agrobacterium vitis."
RL J. Bacteriol. 177:6518-6526(1995).
CC -1- FUNCTION: DEGRADES AN UNIDENTIFIED TOXIC PRODUCT FROM THE FIRST
STEP OF TARTRATE DEGRADATION.
CC -1- CATALYTIC ACTIVITY: D-glycerate + NAD(P)(+) = hydroxypyruvate +
NAD(P)H.
CC -1- PATHWAY: Tartrate degradation; second step.
CC -1- INDUCTION: BY TARTRATE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U25634; AAA68699.1;
CC InterPro; IPR005346; UPF0125.
CC Pfam; PF03658; UPF0125.1.
CC Plasmid; Oxidoreductase; NADP.
SQ SEQUENCE 438 AA; 44730 MW; 38C75BAFFFB0636B CRC64;

Query Match
Best Local Similarity 26.1%; Score 95.5; DB 1; Length 438;
Matches 57; Conservative 28; Mismatches 78; Indels 55; Gaps 13;

QY 52 LLGDELGDGIDGIAICATV--PSVLHELREVTTRYGDPV-----AVLPEGVKTVGPILT 105
Db 202 IISDVPGDPSAEAGTVPADPTTADAAAIARYIDLPESARAVLQ-GNET-----254
QY 106 DHPK--EVGADRIINAARELYGGPAIVVDFGTATTF--DAVSARGEYIGGVIAPIGEI 161
Db 255 --PRAGEVAGEIRLVAPSPISALEAAAAALDAGLCPLILGDALEAGEAREMGVMA-GIAL 311
QY 162 SVEALGVKGAOLRIEVARPSRV1--GKNTVE--AMQSGI-----197
Db 312 SARDKG-----LPAAPAVILSGGESTVSLGNMTEGRGRNTEFLLSLAVALKGASG 363

```

```

QY 198 VYGFAGQVGV--VNMARELADDDPDVTVTIATGGLAP 233
Db 364 IWAIAAGDTGDIGDVEDAAGALV-APDSLINMRDAGIDP 400

RESULT 10
KPYK_MYCTU STANDARD; PRT; 472 AA.
ID KPYK_MYCTU
AC O06134;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pyruvate kinase (EC 2.7.1.40) (PK).
GN PYK OR PYKA OR RV1617 OR MT1653 OR MTCY01B2.09.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsby T., Jagels K., Krogh A., McLean M.A., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Bisbal W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + pyruvate = ADP + phosphoenolpyruvate.
CC -1- COFACTOR: REQUIRES MAGNESIUM AND POTASSIUM.
CC -1- PATHWAY: GLYCOLYSIS; final step.
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PYRUVATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z95554; CAB08894.1;
CC EMBL; AE007029; RAK45923.1;
CC HSP; F14178; IE01.
CC TIGR; MT1653;
CC TuberculList; RV1617;
CC InterPro; IPR001697; Pyruvate_Kinase.
CC Pfam; PF0224; PK; 1.
CC Pfam; PF02887; PK_C; 1.
CC PRINTS; PRO1050; PYRUVTKNASE.
CC PRODOM; PD001009; Pyruvate_kinase; 1.
CC TIGRFAMS; TIGR01064; pyruv_kin; 1.
CC PROSITE; PS00110; PYRUVATE_KINASE; 1.
CC Transferrase; Kinase; Glycolysis; Magnesium; Complete proteome.
KW ACT_SITE 218 218 BY SIMILARITY.
FT ACT_SITE 220 220 MAGNESIUM (BY SIMILARITY).
FT METAL

```

```

DR InterPro: IPR003009; PWM_enzyme.
DR InterPro: IPR002028; Trp_synthaseA.
DR Pfam: PF00290; trp_syntA; 1.
DR ProDom: PD001535; TRP_synthase_alpha; 1.
DR TIGRFAMs: TIGR00262; trpA; 1.
DR PROSITE: PS00167; TRP_SYNTHASE_ALPHA; 1.
KW Tryptophan biosynthesis; Lyase; Complete proteome.
SQ SEQUENCE 279 AA; 29129 MW; 0D83F876B7608FA CRC64;

Query Match 7.1%; Score 95; DB 1; Length 279;
Best Local Similarity 22.7%; Pred. No. 3.6; Gaps 10;
Matches 52; Conservative 31; Mismatches 94; Indels 52; Gaps 10;

QY 19 GEDIVHEWRITDSRRRTADELAVLQ-----LMGMHPLLDGDLGIDGIAICARVPS 72
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 76 GQIVLTKMASFRAGDNTPVLMGYNPIYGYVDRELFKDALASIGDLIVVDLPPE 135

QY 73 VLHELREVTTRYGDVPVLVEFGVKTGVPL---TDHPKEVGADRIINAVAAYELVGGP 129
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 MDEE-----LCIPALKAGINFIRLATPTDDKKRLPKVLQNTSGFVYVSM 180

QY 130 AIYVDFGTATTFDAVSARGEYIGGVITAPGTEISVEALGVKAQLRKEIVARPSVIGKNT 189
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 TGITGSALADTGKVAANRNKIGHTDLP---VCV-GFGVKTA-----EQAR---VIGANA 228

QY 190 VEAQSGIYVGFAGQVDGVNRMAR-----ELADDPDDVTIVATGGIA 232

Db 229 -----DCVVYGTGTA-----IYNVANVLGPRGETDAPEAVATLVSGIA 267

```

RESULT 12

GLK_BACSU	STANDARD;	PRT;	321 AA.
ID	GLK_BACSU	PRT;	321 AA.
AC	P54495;		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Glucokinase (EC 2.7.1.2) (Glucose kinase).		
GN	GLCK.		
OS	Bacillus subtilis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=1423;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=168 / JH642;		
RC	MEDLINE=371124195; PubMed=8969508;		
RA	Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M., Kobayashi Y.;		
RA	"Systematic sequencing of the 283 kb 210 degrees-232 degrees region of the Bacillus subtilis genome containing the skin element and many sporulation genes.";		
RT	Microbiology 142:3103-3111(1996).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=168;		
RC	MEDLINE=36044033; PubMed=9384377;		
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borriest R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschier C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Gim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G., Guiseppi G., Guy B.J., Haga K., Haech J., Harwood C.R., Henaut A., Hilbert B., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Oglwara A., Oudega B., Parf S.H.,		

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*."; 249-256(1997).
RL Nature 390:249-256(1997).
RN [3]
RN CHARACTERIZATION.
RX MEDLINE=98292758; PubMed=9620975;
RA Skarlatos P., Dahl M.K.;
RT "The glucose kinase of *Bacillus subtilis*."; J. Bacteriol. 180:3222-3226(1998).
CC -1- CATALYTIC ACTIVITY: ATP + D-glucose = ADP + D-glucose 6-phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE ROK (NAGC/XLR) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D84432; BAAL2521.1; -
CC EMBL; Z99116; CAB14416.1; -
CC Subtilist; BG11685; glck.
CC InterPro: IPR000600; ROK_family.
CC InterPro: IPR004654; ROK_glcA_fam.
CC Pfam: PF00480; ROK; 1.
CC TIGRFAMS; TIGR00744; ROK_glcA_fam; 1.
CC PROSITE; PS01125; ROK; 1.
CC Transferase; Kinase; Glycolysis; ATP-binding; Complete proteome.
SQ SEQUENCE 321 AA; 33545 MW; DEC2CBC49D801D0 CRC64;

Query Match 7.1%; Score 95; DB 1; Length 321;
Best Local Similarity 22.1%; Pred. No. 4.2;
Matches 64; Conservative 36; Mismatches 97; Indels 92; Gaps 15;
5 IDVGNTHVLGFDGDIH-VRISTDRTADELAVLQGLMGHPLLDGELGDG 63
9 IDLGGTTIKLAFINQYGEIQHKEVPTD-KYGDITV----- 44
64 IATCATVPSVLHLEVRTR--RYYG-----DPAVLVEPGVKTP--ILTDH-PREV 111
45 -TIATIDSKLDELQPKPHIKYIGMGAPGVPDAAAGVYETVNLGWNKYNALKNHLETET 103
112 GADRII---NAVAVELYGG-----PAIVDFGTATFDVARSARGEYIGVIAPIG--- 158
104 GIPAVIENDANIAALGEMKAGDGAKDVLTLGTGVG-GGIANGELHNGAGGGEI 162
159 -----TEISVEALGVKGAQLKTEVARPRSVGVKNVTEAMQSGIV 198
163 GHICISPEGGAPCNGKGTGCTIATATGI---VRIAKGIANAKTKRLKATEQLSARDV 220
199 YGFAGQVD---GVNRMARELADDPDVTVIATGGLA-----PMVLG 237
221 FEARGEDEIALEVVDYVAKHLG-----LVLGNTASSLNPSKIVLG 261

RESULT 13

ID H4_HALN1

ID H4_HALN1

AC Q9H14;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]
DE imidazole-4-carboxamide isomerase (EC 5.3.1.16)
DE (Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide
DE isomerase).
GN HISA OR VNG2294G.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of *Halobacterium* species NRC-1."; Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RL -1- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide = 5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide
CC [(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide.
CC -1- PATHWAY: Histidine biosynthesis; fourth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HISA / HISE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE005113; AAG20407.1; -
CC InterPro: IPR003009; FMN_enzyme.
CC InterPro: IPR000570; His_biosynth.
CC Pfam: PF00977; His_biosynth; 1.
CC TIGRFAMS; TIGR00007; His_biosynth; 1.
CC Isomerase; Histidine biosynthesis; Complete proteome.
SQ SEQUENCE 242 AA; 24649 MW; F51938766AE5FFC1 CRC64;
Query Match 7.0%; Score 93.5; DB 1; Length 242;
Best Local Similarity 24.9%; Pred. No. 3.9;
Matches 61; Conservative 33; Mismatches 80; Indels 71; Gaps 14;
QY 12 TWIGLFDGEDIVHWRISTDSRRTADELAVLQGLMGHPLLDGELGDGIDGTAICATVP 71
DB 7 TVLPATVMDQ-----GQVQLVQGERGTERTYGDPVDAATDWAAGA--- 48
QY 72 SYLHELREVRTRYG-----DPAVLVEPGVKTP--VPILTDPKVEGAD 114
DB 49 EALH-LVLDLGAFAEGARANATAVEDILDATDV-SVQVGGGIRSAEDATALLDR---GVD 102
QY 115 RIINAAVELYGGPAIVDFGTATFDVARSARGEYIGVI---APGTEISVEALGVK 170
DB 103 RVLGTAAIET---PDIVGEIAAA-----YSDGVLVSLDKARDGEVVVEG-WTEG 147
QY 171 AOLRKIEVARPRSVGVKNVTEAMQSGIVY---GFAGQVDGVNRMARELADDPDDVTIA 227
DB 148 TGMDFVAAQRYADLG-----AAGILETVDVVEKQSGVTRDPYRDLVDSV-DIPVIA 199
QY 228 TGGLA 232
DB 200 SGVA 204


```
RESULT 14
PRNA_ECOLI
ID PRNA_ECOLI STANDARD; PRT; 293 AA.
AC P26837; P76680; P76681;
DT 01-DEC-1992 (Rel. 24, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribosomal protein L11 methyltransferase (EC 2.1.1.-).
GN PRNA OR B3259 OR Z4619 OR ECS4131.
OS Escherichia coli, and
OC Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562, 83334;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE-94042890; PubMed-8226664;
RX Vannet A., Plumbidge J.A., Alix J.-H.;
RT "Cotranscription of two genes necessary for ribosomal protein L11
methylation (prma) and pantothenate transport (panf) in Escherichia
coli K-12.";
J. Bacteriol. 175:7178-7188(1993).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
Science 277:1453-1474(1997).
RN [3]
SEQUENCE FROM N.A.
RP STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE-21074935; PubMed-11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
Nature 409:529-533(2001).
RN [4]
SEQUENCE FROM N.A.
RP STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE-21156231; PubMed-11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
DNA Res. 8:11-22(2001).
RN [5]
SEQUENCE OF 140-293 FROM N.A.
RP MEDLINE-93094136; PubMed-1459953;
RA Ball C.A., Osuna R., Ferguson K.C., Johnson R.C.;
RT "Dramatic changes in Fis levels upon nutrient upshift in Escherichia
coli.";
J. Bacteriol. 174:8043-8056(1992).
RN [6]
SEQUENCE OF 1-4 FROM N.A.
RP STRAIN-K12;
RX MEDLINE-90299808; PubMed-2193919;
RA Jackowski S., Alix J.-H.;
RT "Cloning, sequence, and expression of the pantothenate permease
(pant) gene of Escherichia coli.";
J. Bacteriol. 172:3842-3848(1990).
CC -1- FUNCTION: METHYLATES RIBOSOMAL PROTEIN L11.
CC -1- SIMILARITY: TO OTHER METHYLTRANSFERASES.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; S67010; AAB28769.1;
DR EMBL; U18997; AAA58062.1; ALT_FRAME.
DR EMBL; U18997; AAA58063.1; ALT_FRAME.
DR EMBL; AE000404; AAC76291.1;
DR EMBL; AE005553; AAG58387.1;
DR EMBL; AP002564; BAB37554.1;
DR EMBL; M95784; AAA23781.1;
DR EMBL; M30953; NOT_ANNOTATED_CDS.
DR PIR; A47043; A47043.
DR PIR; B49924; B49924.
DR Ecogene; EG11497; prma.
DR InterPro; IPR004498; Prma.
DR InterPro; IPR000051; SAM_bind.
DR TIGRFAMs; TIGR00406; prma; 1.
KW Transferase; Methyltransferase; Complete proteome.
FT CONFLICT 78 L->S (IN REF 1).
SQ SEQUENCE 293 AA; 31877 MW; CF0EDE130245B8FC CRC64;
Query Match 7.0%; Score 93.5; DB 1; Length 293;
Best Local Similarity 24.3%; Pred. No. 4.8;
Matches 71; Conservative 40; Mismatches 82; Indels 99; Gaps 19;
QY 8 GNTHVTVLGFDGEDIVHEWRISTDSRTADSLAVLLQGLMGHPLLG-----DELGD- 59
DB 52 GDT-DVIGLFDLAE-----TDMN---DVVAILEN-----HPLLGAFAHKIQLEDK 93
QY 60 -----GIDGIAICATVPSELHELREVTTRY----YGDVP-----AVLVEPGVKTGVPILT 105
DB 94 DWREWDN-----FHPMRGELWLCPSWRDYPDENAVNVMLDPLGAFCT---G 140
QY 106 DHPKEVGADRIINAVALVYGGPAIVDFDTATTFDAVSARGEYIGGVAPGISVEA 165
DB 141 THPT---TSLCLQWLDSLDTG--KTVIDFGCGSGILAIAL--KLGAAKAIGIDIDPOA 193
QY 166 LGVKGAGQLRIEVA-RPRSVIGKNTVEAMQSGIYVGFAGVDGVVNRMA---RELADDPD 221
DB 194 IQASRDNAERNGVSDRLLELPLKDOPEEMKADV-----VANILAGLPURELA--PL 242
QY 222 DVTVIATGGLAPMVLGESSV-----LGLSGILASQAEVCEAYADSFALDPVVEKEWCRITG 253
DB 243 ISVLPVSGGL-----LGLSGILASQAEVCEAYADSFALDPVVEKEWCRITG 290
RESULT 15
BUK1_CLOAB
ID BUK1_CLOAB STANDARD; PRT; 355 AA.
AC Q45829; Q97EN4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Butyrate kinase 1 (EC 2.7.2.7) (BK1).
GN BUK1 OR BUK OR CAC3075.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE-94063503; PubMed-8244020;
RA Walter K.A., Nair R.V., Cary J.W., Bennett G.N., Papoutsakis E.T.;
RT "Sequence and arrangement of two genes of the butyrate-synthesis
pathway of Clostridium acetobutylicum ATCC 824.";
RL Gene 134:107-111(1993).
RN [2]
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 21:35:06 ; Search time 21.6135 Seconds
(without alignments)
2526.317 Million cell updates/sec

Title: US-09-813-453A-4
Perfect score: 1343
Sequence: 1 MLLTIDVGNTHVIGLFEDGE.....EPWLTMLGLRLYERNVSRM 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
otal number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_viruses.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1343	100.0	265	16 Q9X8N6	Q9x8n6 streptomyc
2	710.5	52.9	274	16 Q9CD56	Q9cd56 mycobacteri
3	698.5	52.0	272	16 O06282	O06282 mycobacteri
4	667	49.7	258	2 Q9F985	Q9f985 bacillus st
5	662	49.3	254	16 Q9KGH5	Q9kgh5 bacillus ha
6	639	47.6	255	16 Q8R7M2	Q8r7m2 thermoanaer
7	636	47.4	259	16 Q8XHL5	Q8xhl5 clostridium
8	596	44.4	259	16 Q8YAC5	Q8yac5 listeria mo
9	596	44.4	273	16 Q97EB4	Q97eb4 clostridium
10	591	44.0	259	16 Q92F54	Q92f54 listeria in
11	517	38.5	261	16 Q9A6Z1	Q9a6z1 caulobacter
12	409.5	30.5	262	16 Q9RX54	Q9rx54 deinococcus
13	373	27.8	256	16 Q8RFE4	Q8rfe4 fusobacteri
14	342.5	25.5	246	16 Q9WZY5	Q9wzy5 thermotoga
15	291.5	21.7	273	16 Q83446	Q83446 treponema p
16	280	20.8	212	2 Q32514	Q32514 desulfovibr

17 181.5 13.5 262 16 051477
18 180.5 13.4 592 16 09JW17
19 177 13.2 592 16 09JXP1
20 174 13.0 257 16 P74045
21 167 12.4 276 16 08YQD7
22 164 12.2 295 16 08Y2M4
23 139 10.3 242 16 09PC14
24 138.5 10.3 248 16 09HWC1
25 128.5 9.6 229 16 067753
26 112 8.3 56 2 P94305
27 107.5 8.0 465 16 Q916C2
28 107 8.0 198 16 0930K7
29 107 8.0 383 16 09CG58
30 105.5 7.9 559 16 09RW79
31 105.5 7.9 1036 2 Q54483
32 102.5 7.6 336 16 Q9WZ57
33 102 7.6 697 5 Q8T8N8
34 102 7.6 755 5 Q8BIC3
35 102 7.6 781 5 Q9GYT0
36 101.5 7.6 293 16 Q8XG12
37 101.5 7.6 368 16 Q9KCL7
38 101 7.5 420 17 Q9YFR8
39 100.5 7.5 522 2 Q9Z692
40 100 7.4 409 2 Q50432
41 100 7.4 1762 2 Q52546
42 100 7.4 1763 2 Q52790
43 100 7.4 3562 2 Q9F829
44 99 7.4 759 2 Q86044
45 99 7.4 1053 16 Q9RK93

ALIGNMENTS

RESULT 1

Q9X8N6 PRELIMINARY; PRT; 265 AA.
AC Q9X8N6;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein SCO3380.
GN SCO3380 OR SCE94.31C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Denapalte D., Eichner A., Cullum J.,
RA Redenbach M., Kieser H.M., Hopwood D.A.;
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdono-Tarraga A.-M., Challis G.L.,
RA Thomson D.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidaigo J., Hornsby T., Howarth S.,

DR Pfam: PF03309; Bvg_acc_factor; 1.
 DR TIGRfams: TIGR00671; baf; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 272 AA; 29304 MW; 5D70B6EF0F09AC8B CRC64;

Query Match 52.0%; Score 698.5; DB 16; Length 272;
 Best Local Similarity 51.5%; Pred. No. 1.6e-42;
 Matches 138; Conservative 53; Mismatches 64; Indels 13; Gaps 3;

QY 1 MLTIDVGNTHVTLGLFDG----EDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDE 56
 DB 1 MLTIDVGNTHVTLGLFDG----EDIVHWRISTDSRRTADELAVLQGLMGHPLGLGDE 56
 QY 57 LGDIDGIAICATVPVSLHRELVTRRYGDPVAVLVEPGVKGVPIILTHPKVEVGADRI 116
 DB 54 DSERLGTALSTVPVSLHRELVTRRYGDPVAVLVEPGVKGVPIILTHPKVEVGADRI 113
 QY 117 INAAVAELVGGPAIVDFGTATFDVARSARGVIGGVIAPIGTEISVEALGVGAQLRKI 176
 DB 114 VNCLEAYDRPKAAIVDFGSSICVDVVSARGFLGGAIAPIGVQVSSDAAARSALURV 173
 QY 177 EVARPSVIGKNTVEAMQSGIYVGFAGQVGVVNRMARELA--DDPDVTVIATGGLAPM 234
 DB 174 ELARPSVIGKNTVEAMQSGIYVGFAGQVGVVNRMARELA--DDPDVTVIATGGLAPM 234
 QY 235 VLGESSVIDHEPWLTLGLRLVYERNV 262
 DB 234 LLPELTVHDYDQHLTLGLRLVYERNV 261

RESULT 4
 ID Q9F985 PRELIMINARY; PRT; 258 AA.
 AC Q9F985;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Putative 32 kDa replication protein.
 OS Bacillus stearothermophilus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Geobacillus.
 OX NCBI_TaxID=1422;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-V;
 RA Vasquez C., Pichuanes S., Saavedra C.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF198621; AAG28531.1;
 DR InterPro: IPR004619; Baf.
 DR Pfam: PF03309; Bvg_acc_factor; 1.
 DR TIGRfams: TIGR00671; baf; 1.
 SQ SEQUENCE 258 AA; 28101 MW; 507B55D695095855 CRC64;

Query Match 49.7%; Score 667; DB 2; Length 258;
 Best Local Similarity 49.6%; Pred. No. 2.8e-40;
 Matches 131; Conservative 57; Mismatches 68; Indels 8; Gaps 5;

QY 1 MLTIDVGNTHVTLGLFDGEDIHWRISTDSRRTADELAVLQGLMGHPLGLGDE 60
 DB 1 MIFVLDVGNTHVTLGLFDGEDIHWRISTDSRRTADELAVLQGLMGHPLGLGDE 60
 QY 61 IDGATCATVPVSLHRELVTRRYGDPVAVLVEPGVKGVPIILTHPKVEVGADRI 120
 DB 57 IDGATCATVPVSLHRELVTRRYGDPVAVLVEPGVKGVPIILTHPKVEVGADRI 115
 QY 121 ANVELVGGPAIVDFGTATFDVARSARGVIGGVIAPIGTEISVEALGVGAQLRKIEVAR 180
 DB 116 AGHLVGSPLIIVDFGTATFDVARSARGVIGGVIAPIGTEISVEALGVGAQLRKIEVAR 175
 QY 181 PRSVIGKNTVEAMQSGIYVGFAGQVGVVNRMARELA--DDPDVTVIATGGLAPMWLG 240
 DB 176 PDDIIGKNTVSAQAGLYGVQVEGIVSRMK---AKSPVPPKVIATGGLASLIASESN 232

QY 241 VIDHEPWLTLGLRLVYERNVSR 264
 DB 233 VIDIVDPFLTLGLKILYKENVDK 256

RESULT 5
 ID Q9KGH5 PRELIMINARY; PRT; 254 AA.
 AC Q9KGH5;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Hypothetical protein BH0086.
 GN BH0086.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001507; BAB03805.1;
 DR InterPro: IPR004619; Baf.
 DR Pfam: PF03309; Bvg_acc_factor; 1.
 DR TIGRfams: TIGR00671; baf; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 254 AA; 27907 MW; 0E3B557BA7DAC176 CRC64;

Query Match 49.3%; Score 662; DB 16; Length 254;
 Best Local Similarity 49.8%; Pred. No. 6.3e-40;
 Matches 132; Conservative 55; Mismatches 64; Indels 14; Gaps 4;

QY 1 MLTIDVGNTHVTLGLFDGEDIHWRISTDSRRTADELAVLQGLMGHPLGLGDE 58
 DB 1 MLTIDVGNTHVTLGLFDGEDIHWRISTDSRRTADELAVLQGLMGHPLGLGDE 58
 QY 59 -DGIDGATCATVPVSLHRELVTRRYGDPVAVLVEPGVKGVPIILTHPKVEVGADRI 117
 DB 54 FQIDIGVIVSSVPPMFMFSLQCKKKYFHTP-MIIGFGIKTGLNIKYNPKVEVGADRI 112
 QY 118 NAAVAELVGGPAIVDFGTATFDVARSARGVIGGVIAPIGTEISVEALGVGAQLRKIE 177
 DB 113 NAAVAELVGGPAIVDFGTATFDVARSARGVIGGVIAPIGTEISVEALGVGAQLRKIE 172
 QY 178 VARPSVIGKNTVEAMQSGIYVGFAGQVGVVNRMARELA--DDPDVTVIATGGLAPMWLG 237
 DB 173 IAKPQVVGTVTIDSMQSGIYVGVQVGVVNRMARELA--DDPDVTVIATGGLAPMWLG 229
 QY 238 ESSVIDHEPWLTLGLRLVYERNV 262
 DB 230 ESETIDVIDSFLTLGLQLYKKNV 254

RESULT 6
 ID Q8R7M2 PRELIMINARY; PRT; 255 AA.
 AC Q8R7M2;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Putative transcriptional regulator, homologs of Bvg accessory
 factor.
 GN TTE2381.
 OS Thermoanaerobacter tengcongensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.


```
QY 121 AAVELYGGPAIVVDFGTATTFDAVSARGEYIGGVIAPIGIEISVEALGVKAOLRKEIVAR 180
DB 116 AASEEYGPVIVVDFGTATTFCTIDESGVTGGAIAPGIMISTEALNRAKLPVRVDIAE 175
QY 181 PRSVIGKNTVEAMQSGIVYGFAGVGVGVNRMARELADDDPDDVTVIATGGLAPMVLGESS 240.
DB 176 SSQIICKSTVSSMQAGIFYGVGQCEGIIAEMKKQSNASP---VVVATGGLARMITEKSS 232
QY 241 VIDEHEPWLTLGMLRLVYERN 261
DB 233 AVDILDPTLTKGLELLYRN 253

RESULT 9
Q97EB4 PRELIMINARY; PRT; 273 AA.
AC Q97EB4;
DT 01-OCT-2001 (Tremblrel. 18, Created)
DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Predicted transcriptional regulator, homolog of Bvg accessory
JE factor.
GN CAC3200.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1486;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=APCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Liu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007815; AAK81136.1; -.
DR InterPro; IPR004619; Baf.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF03309; Bvg_acc_factor; 1.
DR TIGRfams; TIGR00671; baf; 1.
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
RW Complete proteome.
SQ SEQUENCE 273 AA; 30331 MW; DE7B4D5923E72842 CRC64;

Query Match 44.4%; Score 596; DB 16; Length 273;
Best Local Similarity 45.3%; Pred. No. 3.8e-35;
Matches 120; Conservative 58; Mismatches 75; Indels 12; Gaps 4;

QY 1 MLITDVGNTHTVLGLFDGEDIVEHWRISTDSRRTADELAVLLQGLM---GMHPLLGDEL 57
DB 12 VILVLDVGNTHVILVNDKLTAEWRLSTDVLSADEYGIQVNNLFQDQKLDPTL---- 67
QY 58 GDGIDGTAICATVPSVLHRELVTRRYGDPVAVLVEFGVKTGVPILTDHPKEVGADRII 117
DB 68 ---VEGVIISSVVPNIMYSLEHMIRKFKINPLY-VGPGIKTGINIKYDNPKEVGADRI 123
QY 118 NAAVAEYLGPPAIVVDFGTATTFDAVSARGEYIGGVIAPIGIEISVEALGVKAOLRKEIE 177
DB 124 NAAVAHEIYKRLIIDFGTATTCVAVRENGDILGGAICTPGIKVSSEALFEAKLPKVE 183
QY 178 VAPRSVIGKNTVEAMQSGIVYGFAGVGVGVNRMARELADDDP-DVTVIATGGLAPMVL 236
DB 184 LIRPAYAICNTISSIGSIVGYIGQVRYIVRMKELOEKEGEPVAVATGGLAKLIS 243
QY 237 GESSVIDEHEPWLTLGMLRLVYERN 261
DB 244 EEAKNVDVNPFLFLEGLRIIYERN 268

RESULT 10
```

```
Q92F54 PRELIMINARY; PRT; 259 AA.
AC Q92F54;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical protein lin0253.
GN LIN0253.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft U., Kuhn M., Kunst F., Kurapkut G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596164; CAC95486.1; -.
DR Listlist; LIN00253; -.
DR InterPro; IPR004619; Baf.
DR Pfam; PF03309; Bvg_acc_factor; 1.
DR TIGRfams; TIGR00671; baf; 1.
RW Hypothetical protein; Complete proteome.
SQ SEQUENCE 259 AA; 28227 MW; 554B03A0C0EFA64F CRC64;

Query Match 44.0%; Score 591; DB 16; Length 259;
Best Local Similarity 44.8%; Pred. No. 8e-35;
Matches 117; Conservative 55; Mismatches 81; Indels 8; Gaps 3;

QY 1 MLITDVGNTHTVLGLFDGEDIVEHWRISTDSRRTADELAVLLQGLMHPHLLGDELGDG 60
DB 1 MLITDVGNTHTVLGLFDGEDIVEHWRISTDSRRTADELAVLLQGLMHPHLLGDELGDG 60
QY 61 IDGTAICATVPSVLHRELVTRRYGDPVAVLVEFGVKTGVPILTDHPKEVGADRIINAV 120
DB 57 IQGIISSVVPNIMHAMETMCVRYF-NIRPLIVGPGIKTGLNKLVDNPREIGSDRIINAV 115
QY 121 AAVELYGGPAIVVDFGTATTFDAVSARGEYIGGVIAPIGIEISVEALGVKAOLRKEIVAR 180
DB 116 AASEEYGPVIVVDFGTATTFCTIDESGVTGGAIAPGIMISTEALNRAKLPVRVDIAE 175
QY 181 PRSVIGKNTVEAMQSGIVYGFAGVGVGVNRMARELADDDPDDVTVIATGGLAPMVLGESS 240
DB 176 SSQIICKSTVSSMQAGIFYGVGQCEGIIAEMKKQSNASP---VVVATGGLARMITEKSS 232
QY 241 VIDEHEPWLTLGMLRLVYERN 261
DB 233 AVDILDPTLTKGLELLYRN 253

RESULT 11
Q9A6Z1 PRELIMINARY; PRT; 261 AA.
AC Q9A6Z1;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Transcriptional activator, putative, Baf family.
GN CC1935.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
```



```

QY 60 GIDGIAICATVPVSLHRELREVTTRYGVDPAVLVE-----PGVK-TGVPILTDHP 108
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
58 -VDAILSSVVPNIITFFOFARKYF-KVEATIVDEKLPFTFAKINTGT----- 107
QY 109 KEYGADRIINAVALYGGPAIVV-DFGTATTFDAYSARGEYIGGVIAPIEISVEALG 167
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
108 --FGADRIIDITEAMQKYPDKNLVDFDGTATYD-VLKGVGIVGGGILPGIDMSINALY 164
QY 168 VKGAQLRKIEVARPSVIGNTVEAMQSGIVYFAGQVDGVNRMARELADDDPDVTVIA 227
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
165 GNTAKLPVRVFTTSSVLGDTMKQIOAIIFFGAGQIKHIIKINEEL---NEEIFVLA 221
QY 228 TGGAPLVAGESSVIDHEPWLTLMLGLRLVYERN 261
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
222 TGGGLKILSAEIDEIDYDANLSLKGlyTYLKLN 255

RESULT 14
Q9WZY5 PRELIMINARY; PRT; 246 AA.
Q9WZY5
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical protein TM0883.
GN TM0883.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT *Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.*;
RL Nature 399:323-329(1999).
DR EMBL; AE001754; AAD35964.1; -
DR TIGR; TM0883; -
DR InterPro; IPR004619; Baf.
DR Pfam; PF03309; Bvg_acc_factor; 1.
DR TIGRFAMs; TIGR00671; baf; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 246 AA; 27154 MW; 9E0309AD462CF266 CRC64;

Query Match 25.5%; Score 342.5; DB 16; Length 246;
Best Local Similarity 35.7%; Pred. No. 5.2e-17;
Matches 92; Conservative 45; Mismatches 98; Indels 23; Gaps 9;

QY 1 MLITIDVGNTHTVLGLFDGEDIVHEWRISDTSRTADELAVLLOGLMGHPLLDGELGDG 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MYLLVDVGNTHSVFSITEDGKTFRRWRLSTGVFTDEDEL-----FSLHPLLDGAMRE- 53
QY 61 IDGIAICATVPVSLHRELREVTTRYGVDPA-VLVEPG-VKTVGPILTDHPKEVGADRIIN 118
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
54 IKGIGVASVPTQNTVTERSQKYPHFISPIHWKAKNGCVKNV-----KNPSEVGADRVAN 109
QY 119 AVAAVELYGGPAIVVDFGTATTFDAYSARGEYIGGVIAPIEISVEALGVKGAQLRKIEV 178
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
110 VVAPVKEYGKNGIIDMGTTATVDLV-VNGSYEGGAILPGFFMVAHSLFRCTAKLPLEV 168
QY 179 ARPSRVIGKNTVEAMQSGIVYFAGQVDGVNRMARELADDDPDVTVIATGGLAPMV--L 236
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
169 KPADFVYVGKDEENIRLGVNNGSVYALEGIIGRIKEYIGDLP----VVLGGQSKIYKDM 224
QY 237 GESSVIDEHEPWLTLMLGL 254
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
225 IKHEIFDEE---LTIKGV 239

```

```

RESULT 15
O83446 PRELIMINARY; PRT; 273 AA.
ID O83446
AC O83446;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical protein TP0431.
GN TP0431.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artlach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT *Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.*;
RL Science 281:375-388(1998).
DR EMBL; AE001220; AAC65417.1; -
DR TIGR; TP0431; -
DR InterPro; IPR004619; Baf.
DR Pfam; PF03309; Bvg_acc_factor; 1.
DR TIGRFAMs; TIGR00671; baf; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 273 AA; 28472 MW; 439C9C77CB598BC0 CRC64;

Query Match 21.7%; Score 291.5; DB 16; Length 273;
Best Local Similarity 33.1%; Pred. No. 2.7e-13;
Matches 89; Conservative 47; Mismatches 108; Indels 25; Gaps 9;

QY 2 LLTIDVGNTHTVLGLFDGEDIVHEWRISDTSRTADELAVLLOGLMGHPLLDGDE 56
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MLITDVGNTHSVFVGI-QGNGRVCVRELFLAPDARKTQDEVSLLIHALCERAGVGRAS 59
QY 57 LGDGDGIAICATVPVSLHRELREVTTRYGVDPAVLVEPGVKTGVPILTDHP--KEVGAD 114
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
60 LRDAF-----ISSVPEVLTAKTIADAVAQISG-VQPVVFGPWAYEHLFVRIPEPVRABIGTD 114
QY 115 RIINAAVELYGGPAIVVDFGTATTFDAYSARGEYIGGVIAPIEISVEALGVKGAQLR 174
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
115 LVANAAVAVHFRSACVVVDCGTALTFTAVDGTGLIQGVIAIAPGLRTAVOSLHTGTAQLP 174
QY 175 KIEVARPSRVIGKNTVEAMQSGIVYFAGQVDGVNRMARELADDDPDVTVIATGGLAPM 234
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
175 LVPIALPVSIVLGKDTTHAVQAGVVGKTLFVIRAMTAQCKELG---CRCAAVITGGLSRL 231
QY 235 VLGESSVID--EHEPWLTLMLGL-----RLV 257
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
232 F---SSEVDFPPIDAQLTSLGLAHARLV 257

Search completed: June 24, 2003, 21:59:16
Job time : 23.6135 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 22:24:12 ; Search time 14.2826 Seconds
(without alignments)
1947.059 Million cell updates/sec

Title: US-09-813-453A-53

Perfect score: 1303

Sequence: 1 MRLVVDIGNSTTLAIFGWD.....VIDELAVLRGSDLLRMNMP 257

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 41779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 41779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:**

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303	100.0	257	9	US-09-813-453A-53
2	322	24.7	254	9	US-09-813-453A-47
3	319	24.5	258	9	US-09-813-453A-2
4	316	24.3	262	9	US-09-813-453A-45
5	305	23.4	258	9	US-09-813-453A-49
6	301	23.1	255	9	US-09-813-453A-7
7	295	22.6	265	9	US-09-813-453A-4
8	274.5	21.1	272	9	US-09-813-453A-5
9	273.5	21.0	233	9	US-09-813-453A-17
10	271.5	20.8	272	9	US-09-712-363-276
11	265.5	20.4	219	9	US-09-813-453A-57
12	265	20.3	258	9	US-09-813-453A-6
13	261	20.0	260	9	US-09-813-453A-51
14	254.5	19.5	256	9	US-09-813-453A-55
15	253.5	19.5	262	9	US-09-813-453A-8
16	229.5	17.6	250	9	US-09-813-453A-3
17	228.5	17.5	246	9	US-09-813-453A-9
18	227	17.4	241	9	US-09-813-453A-63
19	225	17.3	229	9	US-09-813-453A-12

20	221	17.0	244	9	US-09-813-453A-41	Sequence 41, Appl
21	214	16.4	212	9	US-09-813-453A-59	Sequence 59, Appl
22	197.5	15.2	273	9	US-09-813-453A-10	Sequence 10, Appl
23	166.5	12.8	262	9	US-09-813-453A-11	Sequence 11, Appl
24	162.5	12.5	592	9	US-09-813-453A-43	Sequence 43, Appl
25	161.5	12.4	460	9	US-09-813-453A-39	Sequence 39, Appl
26	159.5	12.2	592	9	US-09-813-453A-22	Sequence 22, Appl
27	157	12.0	248	9	US-09-813-453A-20	Sequence 20, Appl
28	153.5	11.8	249	9	US-09-813-453A-70	Sequence 70, Appl
29	146.5	11.2	257	9	US-09-813-453A-13	Sequence 13, Appl
30	144	11.1	242	9	US-09-813-453A-65	Sequence 65, Appl
31	136.5	10.5	267	9	US-09-813-453A-15	Sequence 15, Appl
32	130.5	10.0	249	9	US-09-813-453A-61	Sequence 61, Appl
33	108	8.3	209	9	US-09-813-453A-21	Sequence 21, Appl
34	105.5	8.1	223	9	US-09-895-913A-74	Sequence 74, Appl
35	105.5	8.1	223	9	US-09-813-453A-14	Sequence 14, Appl
36	105.5	8.1	223	9	US-09-813-453A-67	Sequence 67, Appl
37	98	7.5	787	10	US-09-205-448-8	Sequence 8, Appl
38	84	6.4	574	9	US-09-764-868-774	Sequence 774, Appl
39	83.5	6.4	568	10	US-09-815-242-5140	Sequence 5140, Appl
40	82	6.3	831	10	US-09-815-242-10520	Sequence 10520, Appl
41	80.5	6.2	491	9	US-09-738-626-4429	Sequence 4429, Appl
42	80.5	6.2	491	12	US-10-024-370-3	Sequence 3, Appl
43	80	6.1	5877	9	US-10-142-515-11	Sequence 11, Appl
44	80	6.1	5935	9	US-10-243-243A-8	Sequence 8, Appl
45	79	6.1	991	9	US-10-231-353-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-09-813-453A-53
Sequence 53, Application US/09813453A
Patent No. US20020168681A1
GENERAL INFORMATION:
APPLICANT: Yocum, R. Rogers
APPLICANT: Patterson, Thomas A.
TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF ANTIBIOTICS
FILE REFERENCE: CGZ-001
CURRENT APPLICATION NUMBER: US/09/813,453A
PRIOR FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: US 60/227,860
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 09/567,569
PRIOR FILING DATE: 2000-09-21
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 53
LENGTH: 257
TYPE: PRT
ORGANISM: Chlorobium tepidum
US-09-813-453A-53

Query Match

Best Local Similarity 100.0%; Score 1303; DB 9; Length 257;

Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRLVVDIGNSTTLAIFGWDDEEPSVPSALFADSDSTMTREVFGNMARKHGPQAIAICS	60
Db	1	MRLVVDIGNSTTLAIFGWDDEEPSVPSALFADSDSTMTREVFGNMARKHGPQAIAICS	60
Qy	61	VPSATAVGSALESLSFVPLTTCCKLRPFRLDYATPHFGADRLALCAWRELFSEK	120
Db	61	VPSATAVGSALESLSFVPLTTCCKLRPFRLDYATPHFGADRLALCAWRELFSEK	120
Qy	121	PVIAVDIGTAITFDVLTGVNGYGGGLIMPIDMAGALHSTAOQLPQVRIDRPESILGRS	180
Db	121	PVIAVDIGTAITFDVLTGVNGYGGGLIMPIDMAGALHSTAOQLPQVRIDRPESILGRS	180
Qy	181	TTECIKSGVFWGVKQIGGLVDAIRGDLVRDFGSESTVEIVTGNRIIVPEIGFVSVID	240
Db	181	TTECIKSGVFWGVKQIGGLVDAIRGDLVRDFGSESTVEIVTGNRIIVPEIGFVSVID	240

THIS PAGE BLANK (USPTO)

GenCore version 3.00
Copyright (c) 1993. GenCorp. All rights reserved.

OM protein - protein search, us-09-813-453a-11.rapb
Run on: June 24, 2003, 22:08:46, Search time 11.9453 Seconds
(without alignments)
2373.336 Million cell updates/sec

Title: US-09-813-453A-11
Perfect score: 100.0
Sequence: 1 MNKPLSELIDIGNTSTAFALFKDQVNLFLTKMTNLMRYDEVYFEEFNDFNNKV 262

Scoring table: BLOSUM62
Gap: 10.0, Gapext 0.5

Max hits: 41779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 41779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB_PEP.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB_PEP.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB_PEP.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB_PEP.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB_PEP.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB_PEP.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB_PEP.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB_PEP.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB_PEP.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB_PEP.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB_PEP.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB_PEP.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB_PEP.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1331	100.0	262	9	US-09-813-453A-11
2	282	21.2	256	9	US-09-813-453A-55
3	276.5	20.8	262	9	US-09-813-453A-45
4	275.5	20.7	254	9	US-09-813-453A-47
5	271.5	20.4	258	9	US-09-813-453A-2
6	266.5	20.0	255	9	US-09-813-453A-7
7	262	19.7	250	9	US-09-813-453A-3
8	256.5	19.3	258	9	US-09-813-453A-49
9	231.5	17.4	233	9	US-09-813-453A-17
10	203	15.3	246	9	US-09-813-453A-6
11	197.5	14.8	258	9	US-09-813-453A-5
12	190	14.3	219	9	US-09-813-453A-57
13	189.5	14.2	260	9	US-09-813-453A-51
14	181.5	13.6	263	9	US-09-813-453A-4
15	170	12.8	273	9	US-09-813-453A-10
16	166.5	12.5	257	9	US-09-813-453A-53
17	159	11.9	262	9	US-09-813-453A-8
18	141.5	10.6	229	9	US-09-813-453A-12
19	138	10.4	272	9	US-09-712-363-276

20	138	10.4	272	9	US-09-813-453A-5	Sequence 5, Appl
21	136	10.2	241	9	US-09-813-453A-63	Sequence 63, Appl
22	135	10.1	244	9	US-09-813-453A-41	Sequence 41, Appl
23	121.5	9.1	209	9	US-09-813-453A-21	Sequence 21, Appl
24	121.5	9.1	476	10	US-09-774-414-3	Sequence 3, Appl
25	111	8.3	212	9	US-09-813-453A-59	Sequence 59, Appl
26	100	7.5	592	9	US-09-813-453A-43	Sequence 43, Appl
27	99.5	7.5	410	10	US-09-845-335-2	Sequence 2, Appl
28	96	7.2	592	9	US-09-813-453A-22	Sequence 22, Appl
29	93.5	7.0	242	9	US-09-813-453A-65	Sequence 65, Appl
30	92	6.9	460	9	US-09-813-453A-39	Sequence 39, Appl
31	92	6.9	1161	9	US-10-170-102-4	Sequence 4, Appl
32	90	6.8	257	9	US-09-813-453A-13	Sequence 13, Appl
33	88.5	6.6	223	9	US-09-895-913A-74	Sequence 74, Appl
34	88.5	6.6	223	9	US-09-813-453A-14	Sequence 14, Appl
35	88.5	6.6	223	9	US-09-813-453A-67	Sequence 67, Appl
36	88.5	6.6	370	9	US-10-106-698-6268	Sequence 6268, Ap
37	86.5	6.5	385	10	US-09-815-242-5713	Sequence 5713, Ap
38	86.5	6.5	518	10	US-09-815-242-12473	Sequence 12473, A
39	85.5	6.4	284	9	US-10-284-986-1	Sequence 1, Appl
40	85.5	6.4	284	10	US-09-846-808-1	Sequence 1, Appl
41	84.5	6.3	381	9	US-09-870-759-84	Sequence 84, Appl
42	84.5	6.3	1536	9	US-10-092-880-2	Sequence 2, Appl
43	84	6.3	401	10	US-09-815-242-5506	Sequence 5506, Ap
44	84	6.3	401	10	US-09-815-242-12261	Sequence 12261, A
45	83.5	6.3	248	9	US-09-813-453A-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-09-813-453A-11
; Sequence 11, Application US/09813453A
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; FILE OF INVENTION: ANTIBIOTICS
; FILE REFERENCE: OG2-001
; CURRENT APPLICATION NUMBER: US/09/813,453A
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/227,860
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/667,569
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-813-453A-11

Query Match	100.0%	Score	131	DB 9	Length	262	
Best Local Similarity	100.0%	Pred. No.	5	Seq-ID	115		
Mismatches	262	Conservative	0	Mismatches	0	Gaps	0
QY	1	MNKPLSELIDIGNTSTAFALFKDQVNLFLTKMTNLMRYDEVYFEEFNDFNNKV	60				
DB	1	MNKPLSELIDIGNTSTAFALFKDQVNLFLTKMTNLMRYDEVYFEEFNDFNNKV	60				
QY	61	FISSVVPILNETFKNVIFSFKKIKPLFGDLYDLTFNPNYKSKDFLLGSDVFANLVAAL	120				
DB	61	FISSVVPILNETFKNVIFSFKKIKPLFGDLYDLTFNPNYKSKDFLLGSDVFANLVAAL	120				
QY	121	ENYSENVLVVDLGTACTIFAVSRQDGLGGIINSGLPINFNSLLDNAYLKKFPISTPN	180				
DB	121	ENYSENVLVVDLGTACTIFAVSRQDGLGGIINSGLPINFNSLLDNAYLKKFPISTPN	180				
QY	181	NLLERTTSGSVNSGLFYQYKYLIEGVYRDIKOMYKKFNLIITGGNADLILSLIEFIF	240				

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

protein - protein search, using sw model

Run on: June 24, 2003, 22:08:46 ; Search time 11.1246 seconds
(without alignments)
2373.336 Million cell updates/sec

Title: US-09-813-453A-41

Perfect score: 1229

Sequence: 1 MSFNLIVDQGSACKVAFVR.....ILIHPLVLLGLNRILEYNV 244

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 41779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 41779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PTCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1229	100.0	244	9	US-09-813-453A-41
2	224	18.2	258	9	US-09-813-453A-6
3	221	18.0	257	9	US-09-813-453A-53
4	210.5	17.1	258	9	US-09-813-453A-49
5	208	16.9	255	9	US-09-813-453A-7
6	207.5	16.9	262	9	US-09-813-453A-45
7	207	16.8	258	9	US-09-813-453A-2
8	205.5	16.7	265	9	US-09-813-453A-4
9	201.5	16.4	256	9	US-09-813-453A-55
10	198	16.1	260	9	US-09-813-453A-51
11	197.5	16.1	254	9	US-09-813-453A-47
12	195.5	15.9	248	9	US-09-813-453A-20
13	184	15.0	219	9	US-09-813-453A-70
14	183.5	14.9	241	9	US-09-813-453A-63
15	181.5	14.8	242	9	US-09-813-453A-65
16	181	14.7	249	9	US-09-813-453A-61
17	178	14.5	249	9	US-09-813-453A-70
18	175.5	14.3	233	9	US-09-813-453A-17
19	174	14.2	212	9	US-09-813-453A-59

Sequence 10, Appl
Sequence 3, Appl
Sequence 33, Appl
Sequence 15, Appl
Sequence 9, Appl
Sequence 43, Appl
Sequence 22, Appl
Sequence 276, Appl
Sequence 5, Appl
Sequence 13, Appl
Sequence 11, Appl
Sequence 12, Appl
Sequence 8, Appl
Sequence 74, Appl
Sequence 14, Appl
Sequence 67, Appl
Sequence 21, Appl
Sequence 5530, Appl
Sequence 12355, A
Sequence 12934, A
Sequence 5017, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 19, Appl
Sequence 4265, Appl

20 172.5 14.0 273 9 US-09-813-453A-10
21 167.5 13.6 250 9 US-09-813-453A-3
22 166 13.5 460 9 US-09-813-453A-39
23 163 13.3 267 9 US-09-813-453A-15
24 162 13.2 246 9 US-09-813-453A-9
25 161 13.1 592 9 US-09-813-453A-43
26 156 12.7 592 9 US-09-813-453A-22
27 153 12.4 272 9 US-09-712-363-276
28 152 12.4 272 9 US-09-813-453A-5
29 152 12.4 257 9 US-09-813-453A-13
30 135 11.0 262 9 US-09-813-453A-11
31 132.5 10.8 229 9 US-09-813-453A-12
32 131 10.7 262 9 US-09-813-453A-8
33 129.5 10.5 223 9 US-09-895-913A-74
34 129.5 10.5 223 9 US-09-813-453A-14
35 129.5 10.5 223 9 US-09-813-453A-67
36 126.5 10.3 209 9 US-09-813-453A-21
37 99.5 7.8 468 10 US-09-815-242-5530
38 95.5 7.8 468 10 US-09-815-242-12355
39 95.5 7.8 468 10 US-09-815-242-12934
40 91.5 7.4 783 9 US-09-738-626-5017
41 86 7.0 650 9 US-09-993-241-2
42 86 7.0 650 10 US-09-993-038-2
43 86 7.0 653 9 US-10-206-443-2
44 86 7.0 1140 9 US-09-974-973-19
45 86 7.0 1140 9 US-09-738-626-4265

ALIGNMENTS

RESULT 1

US-09-813-453A-41
; Sequence 41, Application US/09813453A
; Patent No. US20020168681A1

GENERAL INFORMATION:

APPLICANT: Yocum, R. Rogers
APPLICANT: Patterson, Thomas A.
TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
TITLE OF INVENTION: ANTIBIOTICS
FILE REFERENCE: OG2-001
CURRENT APPLICATION NUMBER: US/09/813,453A
CURRENT FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: US 60/227,860
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 09/667,569
PRIOR FILING DATE: 2000-09-21
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 41
LENGTH: 244
TYPE: PRT
ORGANISM: Porphyromonas gingivalis
US-09-813-453A-41

Query Match 100.0%; Score 1229; DB 9; Length 244;
Best Local Similarity 100.0%; Pred. No. 3.3e-108;
Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 MSFNLIVDQGSACKVAFVRNNSIESIFLPKAGQALSHLVAPHKFDKAIYSSVGLPDE 60
Db 1 MSFNLIVDQGSACKVAFVRNNSIESIFLPKAGQALSHLVAPHKFDKAIYSSVGLPDE 60

QY 61 EAEIVRSCAAASLMGTETVPVLRQYDRRTLGADRLAAVVGANSLYFWELLVIDAGT 170
Db 61 EAEIVRSCAAASLMGTETVPVLRQYDRRTLGADRLAAVVGANSLYFWELLVIDAGT 170

QY 121 AITTVRSAGIYLGNNISPGHLRFRKALHFTGRLPLDPSGISPKIAEYGSSTEAIT 180
Db 121 AITTVRSAGIYLGNNISPGHLRFRKALHFTGRLPLDPSGISPKIAEYGSSTEAIT 180

QY 181 AGVTHGLAGEIDRYIDDLHRAKEGRSAVILTGDDANLARIIRSGILIHPLDVLGLGNRL 240
Db 181 AGVTHGLAGEIDRYIDDLHRAKEGRSAVILTGDDANLARIIRSGILIHPLDVLGLGNRL 240

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 22:08:46 ; Search time 12.1733 seconds
(without alignments)
2373.336 Million cell updates/sec

Title: US-09-813-453A-15
Perfect score: 1392
Sequence: 1 MILLIDSGNSRLKVGWFDPPD.....LDSPVLDGLAALAAQGAFTA 267

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41779 seqs, 108206813 residues
Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1392	100.0	267	9	US-09-813-453A-15
2	258.5	18.6	460	9	US-09-813-453A-39
3	256.5	18.4	592	9	US-09-813-453A-22
4	254.5	18.3	592	9	US-09-813-453A-43
5	194.5	14.0	242	9	US-09-813-453A-65
6	178.5	12.8	248	9	US-09-813-453A-20
7	163	11.7	244	9	US-09-813-453A-41
8	163	11.7	249	9	US-09-813-453A-61
9	154.5	11.1	241	9	US-09-813-453A-63
10	152	10.9	246	9	US-09-813-453A-9
11	148	10.6	255	9	US-09-813-453A-7
12	141	10.1	250	9	US-09-813-453A-3
13	140.5	10.1	258	9	US-09-813-453A-49
14	138	9.9	212	9	US-09-813-453A-59
15	138	9.9	262	9	US-09-813-453A-45
16	136.5	9.8	257	9	US-09-813-453A-53
17	134.5	9.7	260	9	US-09-813-453A-51
18	134	9.6	258	9	US-09-813-453A-6
19	134	9.6	265	9	US-09-813-453A-4

20	132	9.5	249	9	US-09-813-453A-70
21	131.5	9.4	254	9	US-09-813-453A-47
22	130.5	9.4	219	9	US-09-813-453A-57
23	129.5	9.3	258	9	US-09-813-453A-2
24	127.5	9.2	256	9	US-09-813-453A-55
25	125	9.0	233	9	US-09-813-453A-17
26	105	7.5	5215	9	US-09-860-846-2
27	105	7.5	5215	9	US-09-988-384B-2
28	105	7.5	5215	9	US-09-836-821-2
29	105	7.5	5215	10	US-09-861-289-2
30	103.5	7.4	262	9	US-09-813-453A-8
31	102.5	7.4	442	9	US-09-464-099A-64
32	102.5	7.4	442	9	US-10-214-766-36
33	102.5	7.4	442	10	US-09-861-696-64
34	97	7.0	257	9	US-09-813-453A-13
35	97	7.0	562	9	US-09-712-363-254
36	95	6.8	408	9	US-09-738-626-5136
37	94.5	6.8	437	9	US-10-145-415-101
38	92.5	6.6	272	9	US-09-712-363-276
39	92.5	6.6	272	9	US-09-813-453A-5
40	90.5	6.5	891	9	US-09-976-059-18
41	89.5	6.4	259	10	US-09-820-893-87
42	89.5	6.4	442	10	US-09-749-728B-11
43	89	6.4	4150	9	US-09-808-880-2
44	88.5	6.4	190	10	US-09-764-864-1071
45	88.5	6.4	1023	9	US-09-893-519A-14

ALIGNMENTS

RESULT 1
US-09-813-453A-15
; Sequence 15, Application US/09813453A
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: ANTIBIOTICS
; FILE REFERENCE: CGZ-001
; CURRENT APPLICATION NUMBER: US/09/813,453A
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/227,860
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/667,569
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-09-813-453A-15

Query Match	100.0%	Score 1392;	DB 9;	Length 267;
Best Local Similarity	100.0%	Pred. No. 3.2e-114;		
Matches 267;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MILLIDSGNSRLKVGWFDPPDPAQAREPAPVAFNDLGLGRWLATLPRRQALGVNV	50	
DB	1	MILLIDSGNSRLKVGWFDPPDPAQAREPAPVAFNDLGLGRWLATLPRRQALGVNV	50	
QY	61	AGLARGAIAATLFRAGGCDIHWLRAQPLAMGLRNGYRNPDLGADRWACMVGVLARQSV	120	
DB	61	AGLARGAIAATLFRAGGCDIHWLRAQPLAMGLRNGYRNPDLGADRWACMVGVLARQSV	120	
QY	121	HPPLLVSFGTATTLDITIGPNVPPGGLILPGPAMRGALAYGTAHLPLADGLVADYD	180	
DB	121	HPPLLVSFGTATTLDITIGPNVPPGGLILPGPAMRGALAYGTAHLPLADGLVADYD	180	
QY	181	THOATAGSIAAQAAGATVROWLAGRQRYGQPEIYVAGGWPVEVROEAKLLAVTGA	240	
DB	181	THOATAGSIAAQAAGATVROWLAGRQRYGQPEIYVAGGWPVEVROEAKLLAVTGA	240	

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

protein search, using sw model

June 24, 2003, 22:24:12 ; Search time 11.7818 Seconds
(without alignments)
1947.059 Million cell updates/sec

Title: US-09-813-453A-59

Perfect score: 1077

Sequence: 1 MTOHFLFDIGNNVKIGIA.....SLNHGFIIFGPAAMTEGVLA 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PTCTUS_PUBCOMB pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1077	100.0	212	9	US-09-813-453A-59
2	379.5	35.2	262	9	US-09-813-453A-45
3	352.5	32.7	258	9	US-09-813-453A-49
4	348.5	32.4	219	9	US-09-813-453A-57
5	338.5	31.4	233	9	US-09-813-453A-17
6	338.5	31.4	255	9	US-09-813-453A-7
7	338.5	31.4	258	9	US-09-813-453A-2
8	330.5	30.7	254	9	US-09-813-453A-47
9	327.5	30.4	250	9	US-09-813-453A-3
10	315.5	29.3	256	9	US-09-813-453A-55
11	307.5	28.6	260	9	US-09-813-453A-51
12	280	26.0	265	9	US-09-813-453A-4
13	271	25.2	262	9	US-09-813-453A-8
14	240.5	22.3	258	9	US-09-813-453A-6
15	221	20.5	246	9	US-09-813-453A-9
16	214	19.9	257	9	US-09-813-453A-53
17	209.5	19.5	272	9	US-09-712-363-276
18	209.5	19.5	272	9	US-09-813-453A-5
19	192.5	17.9	273	9	US-09-813-453A-10

20	180.5	16.8	241	9	US-09-813-453A-63	Sequence 63, Appl
21	174	16.2	244	9	US-09-813-453A-41	Sequence 41, Appl
22	138	12.8	267	9	US-09-813-453A-15	Sequence 15, Appl
23	136.5	12.7	592	9	US-09-813-453A-43	Sequence 43, Appl
24	134	12.4	249	9	US-09-813-453A-70	Sequence 70, Appl
25	133.5	12.4	460	9	US-09-813-453A-39	Sequence 39, Appl
26	132.5	12.3	592	9	US-09-813-453A-22	Sequence 22, Appl
27	120.5	11.2	229	9	US-09-813-453A-12	Sequence 12, Appl
28	114.5	10.6	249	9	US-09-813-453A-61	Sequence 61, Appl
29	111	10.3	262	9	US-09-813-453A-11	Sequence 11, Appl
30	107	9.9	248	9	US-09-813-453A-20	Sequence 20, Appl
31	104	9.7	242	9	US-09-813-453A-65	Sequence 65, Appl
32	102	9.5	223	9	US-09-895-913A-74	Sequence 74, Appl
33	102	9.5	223	9	US-09-813-453A-14	Sequence 14, Appl
34	102	9.5	223	9	US-09-813-453A-67	Sequence 67, Appl
35	92.5	8.6	257	9	US-09-813-453A-13	Sequence 13, Appl
36	88	8.2	275	9	US-10-106-998-6389	Sequence 6389, Ap
37	85	7.9	209	9	US-09-813-453A-21	Sequence 21, Appl
38	82.5	7.7	486	10	US-09-801-368-154	Sequence 154, App
39	82	7.6	250	9	US-10-167-015-2	Sequence 2, Appli
40	81	7.5	376	9	US-09-738-626-6189	Sequence 6189, Ap
41	80.5	7.5	2201	12	US-10-029-907-3	Sequence 3, Appli
42	79.5	7.4	674	9	US-10-086-464-14	Sequence 14, Appl
43	79	7.3	324	10	US-09-815-242-5117	Sequence 5117, Ap
44	77	7.1	4999	9	US-09-976-059-15	Sequence 15, Appl
45	76	7.1	731	9	US-10-086-464-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-813-453A-59
Sequence 59, Application US/09813453A
Patent No. US20020168681A1

GENERAL INFORMATION:
APPLICANT: Yocum, R. Rogers

TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
FILE REFERENCE: CGZ-001

CURRENT APPLICATION NUMBER: US/09/813,453A
CURRENT FILING DATE: 2001-03-20

PRIOR APPLICATION NUMBER: US 60/227,860
PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: 09/667,569
PRIOR FILING DATE: 2000-09-21

NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patentin ver. 2.0

SEQ ID NO 59
LENGTH: 212

TYPE: PRT
ORGANISM: Desulfovibrio vulgaris

US-09-813-453A-59

Query Match 100.0%; Score 1077; DB 9; Length 212;
Best Local Similarity 100.0%; Pred. No. 1e-102;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MTCHFLFDIGNNVKIGIAVETAVLT	SVLPDPGTTDSIGRLLEVLRHAGLGPADV	60
Db	1	MTCHFLFDIGNNVKIGIAVETAVLT	SVLPDPGTTDSIGRLLEVLRHAGLGPADV	60
Qy	61	GACVASSVPGVGNPLIRACERYLYRKL	FAFDIAIPLDNRYRPAEUGADRLVAAYAA	120
Db	61	GACVASSVPGVGNPLIRACERYLYRKL	FAFDIAIPLDNRYRPAEUGADRLVAAYAA	120
Qy	121	RLYPGPRSLVSDVFGTATTFDCV	EGGAYLGGLICPGVLSAGALSRTAKLPRISLEVE	180
Db	121	RLYPGPRSLVSDVFGTATTFDCV	EGGAYLGGLICPGVLSAGALSRTAKLPRISLEVE	180
Qy	181	ESDPVIGRSTTSLNKGFIIFGPAAMTEGV	LAA 212	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

protein - protein search, using sw model

Run on: June 24, 2003, 22:24:12 ; Search time 13.449 Seconds
(without alignments)
1947.059 Million cell updates/sec

Title: US-09-813-453A-65
Perfect score: 1230
Sequence: 1 MNDWFLDGNRSFKASLRE.....HRPTLVLDGLAIWAANV 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues
Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1230	100.0	242	US-09-813-453A-65	Sequence 65, Appl
2	200	16.3	592	US-09-813-453A-43	Sequence 43, Appl
3	199.5	16.2	460	US-09-813-453A-39	Sequence 39, Appl
4	198.5	16.1	592	US-09-813-453A-22	Sequence 22, Appl
5	194.5	15.8	267	US-09-813-453A-15	Sequence 15, Appl
6	193	15.7	248	US-09-813-453A-20	Sequence 20, Appl
7	181.5	14.8	244	US-09-813-453A-41	Sequence 41, Appl
8	168.5	13.7	249	US-09-813-453A-61	Sequence 61, Appl
9	165.5	13.5	249	US-09-813-453A-70	Sequence 70, Appl
10	144	11.7	257	US-09-813-453A-53	Sequence 53, Appl
11	143.5	11.7	219	US-09-813-453A-57	Sequence 57, Appl
12	143.5	11.7	241	US-09-813-453A-63	Sequence 63, Appl
13	140.5	11.4	260	US-09-813-453A-51	Sequence 51, Appl
14	139	11.3	265	US-09-813-453A-4	Sequence 4, Appl
15	134.5	10.9	258	US-09-813-453A-2	Sequence 2, Appl
16	128.5	10.4	262	US-09-813-453A-8	Sequence 8, Appl
17	128	10.4	235	US-09-813-453A-7	Sequence 7, Appl
18	125.5	10.2	258	US-09-813-453A-6	Sequence 6, Appl
19	120	9.8	256	US-09-813-453A-55	Sequence 55, Appl

20	117	9.5	262	9	US-09-813-453A-45	Sequence 45, Appl
21	116	9.4	233	9	US-09-813-453A-17	Sequence 17, Appl
22	115.5	9.4	223	9	US-09-895-913A-74	Sequence 74, Appl
23	115.5	9.4	223	9	US-09-813-453A-14	Sequence 14, Appl
24	115.5	9.4	223	9	US-09-813-453A-67	Sequence 67, Appl
25	115	9.3	258	9	US-09-813-453A-49	Sequence 49, Appl
26	112	9.1	272	9	US-09-712-363-276	Sequence 276, App
27	112	9.1	272	9	US-09-813-453A-5	Sequence 5, Appl
28	108.5	8.8	250	9	US-09-813-453A-3	Sequence 3, Appl
29	107.5	8.7	254	9	US-09-813-453A-47	Sequence 47, Appl
30	105	8.5	498	10	US-09-815-242-11777	Sequence 11777, A
31	104	8.5	212	9	US-09-813-453A-59	Sequence 59, Appl
32	103	8.4	246	9	US-09-813-453A-9	Sequence 9, Appl
33	99	8.0	257	9	US-09-813-453A-13	Sequence 13, Appl
34	95	7.7	229	9	US-09-813-453A-12	Sequence 12, Appl
35	93.5	7.6	262	9	US-09-813-453A-11	Sequence 11, Appl
36	93.5	7.6	323	10	US-09-815-242-13734	Sequence 13734, A
37	93.5	7.6	4999	9	US-09-976-059-15	Sequence 15, Appl
38	89.5	7.3	209	9	US-09-813-453A-21	Sequence 21, Appl
39	88	7.2	273	9	US-09-813-453A-10	Sequence 10, Appl
40	81.5	6.6	653	10	US-09-759-010-2	Sequence 2, Appl
41	81.5	6.6	654	9	US-09-919-039-260	Sequence 260, App
42	81.5	6.6	654	10	US-09-919-172-54	Sequence 54, Appl
43	81.5	6.6	4999	9	US-09-976-059-14	Sequence 14, Appl
44	80.5	6.5	323	10	US-09-815-242-10321	Sequence 10321, A
45	80.5	6.5	342	9	US-09-712-363-210	Sequence 210, App

ALIGNMENTS

RESULT 1
US-09-813-453A-65
; Sequence 65, Application US/09813453A
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: ANTIBIOTICS
; FILE REFERENCE: OGZ-001
; CURRENT APPLICATION NUMBER: US/09/813,453A
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/227,860
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/667,569
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
; US-09-813-453A-65

Query Match	100.0%	Score 1230;	DB 9;	Length 242;
Best Local Similarity	100.0%	Pred. No. 7.3e-107;		
Matches 242;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MNDWFLDGNRSFKASLREGVIGPVTVLPYLTETMDAFALQELPRGRVAYLASVAAPAI	60	
DB	1	MNDWFLDGNRSFKASLREGVIGPVTVLPYLTETMDAFALQELPRGRVAYLASVAAPAI	60	
QY	61	TTTHVLEVLKIHFEQVQVAAATVAACAGVRIAYAHPERGVDREFALLGSGGNNLVVGVG	120	
DB	61	TTTHVLEVLKIHFEQVQVAAATVAACAGVRIAYAHPERGVDREFALLGSGGNNLVVGVG	120	
QY	121	TALTIDLLAANGCHIGGRISASPTLMRQALHARAQPLSGGNYLEFAEDTALVSGCN	180	
DB	121	TALTIDLLAANGCHIGGRISASPTLMRQALHARAQPLSGGNYLEFAEDTALVSGCN	180	
QY	181	GAVALTERSLYEAHORLDQSVRLHGGGVASLLPWLGVDVHVRPTLVLDGLATAWAAVA	240	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

protein - protein search, using sw model

Run on: June 24, 2003, 22:24:12 ; Search time 12.3931 Seconds
(without alignments)
1947.059 Million cell updates/sec

Title: US-09-813-453A-67
Perfect score: 1153
Sequence: 1 MPARQSFDTLKNLVLCIDGN.....RLVFDGMEALKAGILECK 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues
Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB_PEP.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB_PEP.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB_PEP.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB_PEP.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB_PEP.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB_PEP.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB_PEP.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB_PEP.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB_PEP.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB_PEP.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB_PEP.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB_PEP.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB_PEP.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1153	100.0	223	9	US-09-895-913A-74
2	1153	100.0	223	9	US-09-813-453A-14
3	1153	100.0	223	9	US-09-813-453A-67
4	391.5	34.0	209	9	US-09-813-453A-21
5	165	14.3	246	9	US-09-813-453A-9
6	159.5	13.8	248	9	US-09-813-453A-20
7	148	12.8	249	9	US-09-813-453A-70
8	144.5	12.5	241	9	US-09-813-453A-63
9	144.5	12.5	249	9	US-09-813-453A-13
10	141.5	12.3	257	9	US-09-813-453A-12
11	138.5	12.0	229	9	US-09-813-453A-41
12	129.5	11.2	244	9	US-09-813-453A-22
13	116.5	10.1	592	9	US-09-813-453A-43
14	116.5	10.1	592	9	US-09-813-453A-43
15	116	10.1	460	9	US-09-813-453A-39
16	115.5	10.0	242	9	US-09-813-453A-17
17	109.5	9.5	233	9	US-09-813-453A-2
18	109.5	9.5	258	9	US-09-813-453A-2
19	108	9.4	219	9	US-09-813-453A-57

20	105.5	9.2	257	9	US-09-813-453A-53
21	102.5	8.9	250	9	US-09-813-453A-3
22	102	8.8	212	9	US-09-813-453A-59
23	101	8.8	256	9	US-09-813-453A-55
24	100.5	8.7	255	9	US-09-813-453A-7
25	98.5	8.5	262	9	US-09-813-453A-45
26	91	7.9	254	9	US-09-813-453A-47
27	89.5	7.8	272	9	US-09-712-363-276
28	89.5	7.8	272	9	US-09-813-453A-11
29	88.5	7.7	262	9	US-09-813-453A-49
30	87.5	7.6	258	9	US-09-813-453A-4
31	87	7.5	947	10	US-09-801-574-4
32	84	7.3	764	10	US-09-815-242-11556
33	82.5	7.2	265	9	US-09-813-453A-4
34	82	7.1	688	10	US-09-815-242-11416
35	81.5	7.1	273	9	US-09-813-453A-10
36	80	6.9	260	9	US-09-813-453A-51
37	79	6.9	310	10	US-09-815-242-10283
38	79	6.9	375	9	US-09-981-353-113
39	79	6.9	2799	9	US-10-151-736-4
40	78	6.8	770	9	US-10-153-668-7
41	77.5	6.7	258	9	US-09-813-453A-6
42	77	6.7	692	9	US-10-101-464A-897
43	76	6.6	434	9	US-10-153-668-454
44	76	6.6	439	9	US-09-374-046A-44
45	76	6.6	653	9	US-10-174-590-402

ALIGNMENTS

RESULT 1
US-09-895-913A-74
; Sequence 74, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Comen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1 Helicobacter Polypeptides in t
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/981,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-74

Query Match	100.0%	Score 1153	DB 9	Length 223
Best Local Similarity	100.0%	Pred. No. 1.8e-107		
Matches 223	Conservative 0	Mismatches 0	Indels 0	Gaps
QY	1	MPARQSFDTLKNLVLCIDGNTRHFAQNTQLFSSAKEDKRLGKIQKEIFYTSVNEENEP		
Db	1	MPARQSFDTLKNLVLCIDGNTRHFAQNTQLFSSAKEDKRLGKIQKEIFYTSVNEENEP		
QY	61	LLMCPYNAKNIAGFFHLETDYVGLIDRQMACIAVNNVGVVDAGSATTIDILKEGKHLG		
Db	61	LLMCPYNAKNIAGFFHLETDYVGLIDRQMACIAVNNVGVVDAGSATTIDILKEGKHLG		
QY	121	CILPGLAQYTHAYKSAKILEQPFKALDSLEVLPKSTRDAVNYGMVLSVIACIQHLAKNQ		
Db	121	CILPGLAQYTHAYKSAKILEQPFKALDSLEVLPKSTRDAVNYGMVLSVIACIQHLAKNQ		

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003: Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 22:08:46 ; Search time 10.1672 Seconds
(without alignments)
2373.336 Million cell updates/sec

Title: US-09-813-453A-14
Perfect score: 1153
Sequence: 1 MPARQSFDTLKNLVCLDIGN.....RLVPDGMELAKKAGILECK 223

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PCTUS_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
15: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1153	100.0	223	9	US-09-895-913A-74
2	1153	100.0	223	9	US-09-813-453A-14
3	1153	100.0	223	9	US-09-813-453A-67
4	391.5	34.0	209	9	US-09-813-453A-21
5	165	13.8	246	9	US-09-813-453A-9
6	159.5	13.8	248	9	US-09-813-453A-20
7	148	12.8	249	9	US-09-813-453A-70
8	144.5	12.5	241	9	US-09-813-453A-63
9	144.5	12.5	249	9	US-09-813-453A-61
10	141.5	12.3	257	9	US-09-813-453A-13
11	138.5	12.0	229	9	US-09-813-453A-12
12	129.5	11.2	244	9	US-09-813-453A-41
13	116.5	10.1	592	9	US-09-813-453A-22
14	116.5	10.1	592	9	US-09-813-453A-43
15	116	10.1	460	9	US-09-813-453A-39
16	115.5	10.0	242	9	US-09-813-453A-65
17	109.5	9.5	233	9	US-09-813-453A-17
18	109.5	9.5	258	9	US-09-813-453A-2
19	108	9.4	219	9	US-09-813-453A-57

20	105.5	9.2	257	9	US-09-813-453A-53
21	102.5	8.9	250	9	US-09-813-453A-3
22	102	8.8	212	9	US-09-813-453A-59
23	101	8.8	256	9	US-09-813-453A-55
24	100.5	8.7	255	9	US-09-813-453A-7
25	98.5	8.5	262	9	US-09-813-453A-45
26	91	7.9	254	9	US-09-813-453A-47
27	89.5	7.8	272	9	US-09-712-363-276
28	89.5	7.8	272	9	US-09-813-453A-5
29	88.5	7.7	262	9	US-09-813-453A-11
30	87.5	7.6	258	9	US-09-813-453A-49
31	87	7.5	947	10	US-09-801-574-4
32	84	7.3	764	10	US-09-815-242-11556
33	82.5	7.2	265	9	US-09-813-453A-4
34	82	7.1	688	10	US-09-815-242-11416
35	81.5	7.1	273	9	US-09-813-453A-10
36	80	6.9	260	9	US-09-813-453A-51
37	79	6.9	310	10	US-09-815-242-10283
38	79	6.9	375	9	US-09-981-353-113
39	79	6.9	2799	9	US-10-151-736-4
40	78	6.8	770	9	US-10-153-668-7
41	77.5	6.7	258	9	US-09-813-453A-6
42	77	6.7	692	9	US-10-101-464A-897
43	76	6.6	434	9	US-10-153-668-454
44	76	6.6	439	9	US-09-374-046A-44
45	76	6.6	653	9	US-10-174-590-402

ALIGNMENTS

RESULT 1
US-09-895-913A-74
; Sequence 74, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1 Helicobacter P.
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-74

Query Match	100.0%	Score 1153;	DB 9;	Length 223;
Best Local Similarity	100.0%	Pred. No. 1.8e-107;		
Matches 223;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MPARQSFDTLKNLVCLDIGNTRIHFPAQYQLFSSAKEDLKRIGQKEIFYTSVNEENKA	60	
DB	1	MPARQSFDTLKNLVCLDIGNTRIHFPAQYQLFSSAKEDLKRIGQKEIFYTSVNEENKA	60	
QY	61	LLNCYPNAKNIAGFFHLETDYVGLGIDROMACLAANNVGVVDAGSAITDILKEGKHLGG	120	
DB	61	LLNCYPNAKNIAGFFHLETDYVGLGIDROMACLAANNVGVVDAGSAITDILKEGKHLGG	120	
QY	121	CILPGLAQYTHAYKKSAILKQFPKALDSLEVLFPKSTRDAVNYGMVLSVIAICIOHLAKNQ	180	
DB	121	CILPGLAQYTHAYKKSAILKQFPKALDSLEVLFPKSTRDAVNYGMVLSVIAICIOHLAKNQ	180	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

c) 1993 - 2003 Comput

protein - protein search, using sw model

Run on: June 24, 2003, 22:08:46 ; Search time 20.9726 Seconds
(without alignments)
2373.336 Million cell updates/sec

Title: US-09-813-453A-39
 Perfect score: 2384
 Sequence: 1 MGECIMSFQWAFRRPOYEL.....LVHICLLNLIAEGGSESHA 460

Scoring table: BLOSUM62 ;
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues
Total number of hits satisfying chosen parameters: 417779

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : Published_Applications_AA:*
1: /cgn2 6/ptodata/2/pubpaa/US08 NEW PUB. pep:
```

2: /cgn2_6/ptodata/2/pubppaa/FCI_USR1_NEW_PUB.pbp.
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pbp.
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pbp.
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pbp.
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pbp.
7: /cgn2_6/ptodata/2/pubppaa/FCI_USR1_NEW_PUB.pbp.
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pbp.
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pbp.
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pbp.
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pbp.
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pbp.
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pbp.
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pbp.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is defined by analogy of the total score distribution.

COMPARTES

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	2384	100.0	460	9	US-09-813-453A-39	Sequence 39, Appl
2	2223	93.2	592	9	US-09-813-453A-22	Sequence 22, Appl
3	2197	92.2	592	9	US-09-813-453A-43	Sequence 43, Appl
4	258.5	10.8	267	9	US-09-813-453A-15	Sequence 15, Appl
5	232.5	9.8	248	9	US-09-813-453A-20	Sequence 20, Appl
6	232	9.7	249	9	US-09-813-453A-61	Sequence 61, Appl
7	204	8.6	302	9	US-10-260-877-110	Sequence 110, App
8	199.5	8.4	242	9	US-09-813-453A-65	Sequence 65, Appl
9	185	7.8	260	9	US-09-813-453A-51	Sequence 51, Appl
10	179	7.5	265	9	US-09-813-453A-4	Sequence 4, Appl
11	177	7.4	249	9	US-09-813-453A-70	Sequence 70, Appl
12	169	7.1	250	9	US-09-813-453A-3	Sequence 3, Appl
13	168	7.0	262	9	US-09-813-453A-45	Sequence 45, Appl
14	166	7.0	244	9	US-09-813-453A-41	Sequence 41, Appl
15	161.5	6.8	237	9	US-09-813-453A-53	Sequence 53, Appl
16	153	6.4	219	9	US-09-813-453A-57	Sequence 57, Appl
17	153	6.4	288	9	US-09-738-626-4286	Sequence 4286, Ap
18	152.5	6.4	246	9	US-09-813-453A-9	Sequence 9, Appl
19	151	6.3	258	9	US-09-813-453A-2	Sequence 2, Appl

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

Protein - protein search, using sw model

Run on: June 24, 2003, 22:08:46 ; Search time 26.9909 Seconds
(without alignments)
2373.336 Million cell updates/sec

Title: US-09-813-453A-22
Perfect score: 3080
Sequence: 1 MVLKPSHWRLAEADGLP.....LVHGLNLIAARGGESEHT 592

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41779 seqs, 108206813 residues 417779

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3080	100.0	592	9 US-09-813-453A-22	Sequence 22, Appl
2	3024	98.2	592	9 US-09-813-453A-43	Sequence 43, Appl
3	2223	72.2	460	9 US-09-813-453A-39	Sequence 39, Appl
4	256.5	8.3	267	9 US-09-813-453A-15	Sequence 15, Appl
5	256	8.3	302	9 US-10-260-877-110	Sequence 110, App
6	243	7.9	249	9 US-09-813-453A-61	Sequence 61, Appl
7	231.5	7.5	248	9 US-09-813-453A-20	Sequence 20, Appl
8	198.5	6.4	242	9 US-09-813-453A-65	Sequence 65, Appl
9	187	6.1	249	9 US-09-813-453A-70	Sequence 70, Appl
10	183.5	6.0	288	9 US-09-738-626-4286	Sequence 4286, Ap
11	180.5	5.9	265	9 US-09-813-453A-4	Sequence 4, Appl1
12	179	5.8	311	10 US-09-859-270-2	Sequence 2, Appl1
13	175	5.7	260	9 US-09-813-453A-51	Sequence 51, Appl
14	168	5.5	250	9 US-09-813-453A-3	Sequence 3, Appl1
15	168	5.5	262	9 US-09-813-453A-45	Sequence 45, Appl
16	159.5	5.2	257	9 US-09-813-453A-53	Sequence 53, Appl
17	156	5.1	244	9 US-09-813-453A-41	Sequence 41, Appl
18	152	4.9	219	9 US-09-813-453A-57	Sequence 57, Appl
19	152	4.9	272	9 US-09-712-363-276	Sequence 276, App

20	152	4.9	272	9 US-09-813-453A-5	Sequence 5, Appl
21	150.5	4.9	246	9 US-09-813-453A-9	Sequence 9, Appl
22	150	4.9	258	9 US-09-813-453A-2	Sequence 2, Appl
23	147.5	4.8	273	9 US-09-813-453A-10	Sequence 10, A
24	146	4.7	258	9 US-09-813-453A-49	Sequence 49, A
25	141	4.6	255	9 US-09-813-453A-7	Sequence 7, A
26	141	4.6	257	9 US-09-813-453A-13	Sequence 13, A
27	140.5	4.6	256	9 US-09-813-453A-55	Sequence 55, A
28	140	4.5	258	9 US-09-813-453A-6	Sequence 6, A0
29	137	4.4	254	9 US-09-813-453A-17	Sequence 17, A
30	133	4.3	233	9 US-09-813-453A-47	Sequence 47, A
31	132.5	4.3	212	9 US-09-813-453A-59	Sequence 59, A
32	116.5	3.8	223	9 US-09-895-913A-74	Sequence 74, A
33	116.5	3.8	223	9 US-09-813-453A-14	Sequence 14, A
34	116.5	3.8	223	9 US-09-813-453A-67	Sequence 67, A
35	115	3.7	3816	9 US-09-808-880-3	Sequence 3, A2
36	112.5	3.7	241	9 US-09-813-453A-63	Sequence 63, A2
37	112	3.6	472	10 US-09-815-242-13727	Sequence 13727
38	110	3.6	340	10 US-09-815-242-11743	Sequence 11743
39	108.5	3.5	286	9 US-09-769-787-100	Sequence 100
40	108	3.5	229	9 US-09-813-453A-12	Sequence 12, A
41	106.5	3.5	335	10 US-09-815-242-5138	Sequence 5138
42	105	3.4	1222	10 US-09-137-531-15	Sequence 15, A
43	105	3.4	1252	10 US-09-137-531-9	Sequence 9, A
44	104.5	3.4	4613	9 US-09-860-846-31	Sequence 31, A
45	104.5	3.4	4613	9 US-09-988-384B-31	Sequence 31, A

ALIGNMENTS

RESULT 1
US-09-813-453A-22
; Sequence 22, Application US/09813453A
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Vocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION
; TITLE OF INVENTION: ANTIBIOTICS
; FILE REFERENCE: OG2-001
; CURRENT FILING DATE: 2001-03-20
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/227,860
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/667,569
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-813-453A-22

Query Match	100.0%	Score	3080;	DB	9;	Length	592;
Best Local Similarity	100.0%	Pred. No.	3.7e-251;				
Matches	592;	Conservative	0;	Mismatches	0;	Indels	0;
Gap							
QY	1	MTVLKPSHWRLAEADGLPQHV	SOLARVADMKPQOLNGFWOOMP	PAHIRGLLRQHDGY			
Db	1	MTVLKPSHWRLAEADGLPQHV	SOLARVADMKPQOLNGFWOOMP	PAHIRGLLRQHDGY			
QY	61	LVRPLAVFDAGELRELGERSGFT	ALKHECASSNDEILELARIAPDKA	HKHTICVTHLC			
Db	61	LVRPLAVFDAGELRELGERSGFT	ALKHECASSNDEILELARIAPDKA	HKHTICVTHLC			
QY	121	GGRGGRKSHRLGECIMFSFGWV	FDPRPOYELGSLSPVAACRRAL	SLRGLKTKQY			
Db	121	GGRGGRKSHRLGECIMFSFGWV	FDPRPOYELGSLSPVAACRRAL	SLRGLKTKQY			
QY	181	DLWVGDKLGGILLIETVRTGK	TVAVVIGINFLVLPKEVENA	SWCFLPFAERGN			

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

protein search, using sw model

June 24, 2003, 22:08:46 ; Search time 26.9909 Seconds
(without alignments)

2373.336 Million cell updates/sec

Title: US-09-813-453A-43

Perfect score: 3078

Sequence: 1 MTVLKLSHRWVLAELADGLP.....LVIYGLNMTAAGREVEHI 592

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3078	100.0	592	9	US-09-813-453A-43
2	3024	98.2	592	9	US-09-813-453A-22
3	2197	71.4	460	9	US-09-813-453A-39
4	259	8.4	302	9	US-10-260-877-110
5	254.5	8.3	267	9	US-09-813-453A-15
6	245	8.0	249	9	US-09-813-453A-61
7	234.5	7.6	248	9	US-09-813-453A-20
8	200	6.5	242	9	US-09-813-453A-65
9	194	6.3	311	10	US-09-859-270-2
10	191	6.2	249	9	US-09-813-453A-70
11	184.5	6.0	288	9	US-09-738-626-4286
12	177	5.8	265	9	US-09-813-453A-4
13	172	5.6	260	9	US-09-813-453A-51
14	169.5	5.5	262	9	US-09-813-453A-45
15	168	5.5	250	9	US-09-813-453A-3
16	162.5	5.3	257	9	US-09-813-453A-53
17	161	5.2	244	9	US-09-813-453A-41
18	150	4.9	219	9	US-09-813-453A-57
19	150	4.9	258	9	US-09-813-453A-2

20	147.5	4.8	272	9	US-09-712-363-276	Sequence 276, Appl
21	147.5	4.8	272	9	US-09-813-453A-5	Sequence 5, Appl
22	147.5	4.8	273	9	US-09-813-453A-10	Sequence 10, Appl
23	146	4.7	248	9	US-09-813-453A-9	Sequence 9, Appl
24	146	4.7	258	9	US-09-813-453A-49	Sequence 49, Appl
25	141	4.6	227	9	US-09-813-453A-13	Sequence 13, Appl
26	140.5	4.6	256	9	US-09-813-453A-55	Sequence 55, Appl
27	139.5	4.5	255	9	US-09-813-453A-7	Sequence 7, Appl
28	136.5	4.4	212	9	US-09-813-453A-59	Sequence 59, Appl
29	136	4.4	254	9	US-09-813-453A-47	Sequence 47, Appl
30	134	4.4	258	9	US-09-813-453A-6	Sequence 6, Appl
31	133	4.3	223	9	US-09-813-453A-17	Sequence 17, Appl
32	116.5	3.8	223	9	US-09-895-913A-74	Sequence 74, Appl
33	116.5	3.8	223	9	US-09-813-453A-14	Sequence 14, Appl
34	116.5	3.8	223	9	US-09-813-453A-67	Sequence 67, Appl
35	115.5	3.8	241	9	US-09-813-453A-63	Sequence 63, Appl
36	114.5	3.7	3816	9	US-09-808-880-3	Sequence 3, Appl
37	105.5	3.4	4613	9	US-09-860-846-31	Sequence 31, Appl
38	105.5	3.4	4613	9	US-09-988-384B-31	Sequence 31, Appl
39	105.5	3.4	4613	9	US-09-836-821-31	Sequence 31, Appl
40	105.5	3.4	4613	10	US-09-861-289-31	Sequence 31, Appl
41	105.5	3.4	11877	9	US-09-860-846-6	Sequence 6, Appl
42	105.5	3.4	11877	9	US-09-836-821-6	Sequence 6, Appl
43	105.5	3.4	11877	10	US-09-861-289-6	Sequence 6, Appl
44	105.5	3.4	12199	9	US-09-988-384B-6	Sequence 6, Appl
45	104.5	3.4	340	10	US-09-815-242-11743	Sequence 11743, A

ALIGNMENTS

RESULT 1

US-09-813-453A-43
Sequence 43, Application US/09813453A
Patent No. US20020168681A1

GENERAL INFORMATION:
APPLICANT: Iocum, R. Rogers

APPLICANT: Patterson, Thomas A.

TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF

TITLE OF INVENTION: ANTIBIOTICS

FILE REFERENCE: OGZ-001

CURRENT APPLICATION NUMBER: US/09/813,453A
CURRENT FILING DATE: 2001-03-20

PRIOR APPLICATION NUMBER: US 60/227,860
PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: 09/667,569
PRIOR FILING DATE: 2000-09-21

NUMBER OF SEQ IDS: 77

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 43

LENGTH: 592

TYPE: PRT

ORGANISM: Neisseria meningitidis

US-09-813-453A-43

Query Match

Best Local Similarity 100.0%; Score 3078; DB 9; Length 592;

Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MTVLKLSHRWVLAELADGLPQHVLSOLARMADMKPOOLNGFWOOMPAHIRGLRQHDGYWR 60

1 MTVLKLSHRWVLAELADGLPQHVLSOLARMADMKPOOLNGFWOOMPAHIRGLRQHDGYWR 60

61 LVRPLAVFDAGRLGERSGFOTALKHECASSNDEILELARIAPDKAHTTCVTHLOSK 120

61 LVRPLAVFDAGRLGERSGFOTALKHECASSNDEILELARIAPDKAHTTCVTHLOSK 120

121 GRGQRKWSRLGECLEMFSGWVDFDRPOYELGSLSPVAACRRLSRGLDVLQKPN 180

121 GRGQRKWSRLGECLEMFSGWVDFDRPOYELGSLSPVAACRRLSRGLDVLQKPN 180

181 DLVVRDLKGLITETVTGKTVAVVIGINFVLPKEVENAASVOSLFOFQTSRRGNADA 240

181 DLVVRDLKGLITETVTGKTVAVVIGINFVLPKEVENAASVOSLFOFQTSRRGNADA 240

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright. (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 22:24:12 ; Search time 14.5605 Seconds
(without alignments)
1947.059 Million cell updates/sec

Title: US-09-813-453A-45
Perfect score: 1341
Sequence: 1 MIFVLDVGNNTAVLGVPFEG.....LKGLYMLYERNALQHEKGE 262

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues
Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1341	100.0	262	9	US-09-813-453A-45
2	1057.5	78.9	258	9	US-09-813-453A-49
3	1034	77.1	258	9	US-09-813-453A-2
4	960	71.6	254	9	US-09-813-453A-47
5	901	67.2	233	9	US-09-813-453A-17
6	802.5	59.8	256	9	US-09-813-453A-55
7	783	58.4	255	9	US-09-813-453A-7
8	665.5	49.6	250	9	US-09-813-453A-3
9	638	47.6	265	9	US-09-813-453A-4
10	566	42.2	260	9	US-09-813-453A-4
11	536	40.0	258	9	US-09-813-453A-51
12	527	39.3	219	9	US-09-813-453A-57
13	471.5	35.2	272	9	US-09-813-453A-5
14	467.5	34.9	272	9	US-09-712-363-276
15	405.5	30.2	262	9	US-09-813-453A-8
16	399.5	29.8	246	9	US-09-813-453A-9
17	379.5	28.3	212	9	US-09-813-453A-59
18	342.5	25.5	273	9	US-09-813-453A-10
19	316	23.6	257	9	US-09-813-453A-53

20	276.5	20.6	262	9	US-09-813-453A-11
21	207.5	15.5	244	9	US-09-813-453A-61
22	190.5	14.2	241	9	US-09-813-453A-53
23	169.5	12.6	592	9	US-09-813-453A-83
24	168	12.5	460	9	US-09-813-453A-39
25	168	12.5	592	9	US-09-813-453A-22
26	167.5	12.5	229	9	US-09-813-453A-12
27	156.5	11.7	257	9	US-09-813-453A-13
28	142	10.6	249	9	US-09-813-453A-61
29	138	10.3	267	9	US-09-813-453A-15
30	137	10.2	249	9	US-09-813-453A-70
31	136	10.1	248	9	US-09-813-453A-20
32	117	8.7	242	9	US-09-813-453A-65
33	113	8.4	209	9	US-09-813-453A-21
34	98.5	7.3	223	9	US-09-895-913A-74
35	98.5	7.3	223	9	US-09-813-453A-14
36	98.5	7.3	223	9	US-09-813-453A-67
37	97	7.2	509	10	US-09-815-242-5796
38	97	7.2	517	10	US-09-815-242-12837
39	90	6.7	294	10	US-09-815-242-13475
40	88.5	6.6	1019	1	US-08-834-705-18
41	86	6.4	357	9	US-10-012-895-826
42	86	6.4	357	9	US-09-895-793-826
43	86	6.4	357	9	US-09-895-814-826
44	86	6.4	357	10	US-09-759-143-826
45	86	6.4	357	10	US-09-780-669-826

ALIGNMENTS

RESULT 1
US-09-813-453A-45
; Sequence 45, Application US/09813453A
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; FILE REFERENCE: OG2-001
; CURRENT APPLICATION NUMBER: US/09/813,453A
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/227,860
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/667,569
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-813-453A-45

Query Match	100.0%	Score	1341	DB	9	Length	262
Best Local Similarity	100.0%	Pred. No.	3.9e+121	Mismatches	0	Indels	0
Matches	262	Conservative	0				
QY	1	MIFVLDVGNNTAVLGVPFEGELRQHRMETDRHKTDEYGMVLKQLLEHGLSPEDYVKGI	60				
Db	1	MIFVLDVGNNTAVLGVPFEGELRQHRMETDRHKTDEYGMVLKQLLEHGLSPEDYVKGI	60				
QY	61	IVSSVVPPIPMFALEMCERKFKIKPLVVGPGKTKLNKIKYENPREVGADRVNVAAGIHL	120				
Db	61	IVSSVVPPIPMFALEMCERKFKIKPLVVGPGKTKLNKIKYENPREVGADRVNVAAGIHL	120				
QY	121	YGSPLIIVDFGTATTTCYINEEKHYMGVITPGIMISAAALYKRAALRIETTKSSVV	180				
Db	121	YGSPLIIVDFGTATTTCYINEEKHYMGVITPGIMISAAALYKRAALRIETTKSSVV	180				
QY	181	GNNTYSAMOSGLLYCVGOVEGIVKRMKEAKOEKVIATGGLAKLTSTESNVIVYVQCF	240				

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Protein - protein search, using sw model

Run on: June 24, 2003, 22:24:12 ; Search time 14.1159 Seconds
(without alignments)
1947.059 Million cell updates/sec

Title: US-09-813-453A-47
Perfect score: 1293
Sequence: 1 MLIVDVGNTVGLGYQDE.....DVIDSELTGLQIYKKNV 254

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues
Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications_AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1293	100.0	254	9	US-09-813-453A-47
2	987	76.3	258	9	US-09-813-453A-2
3	975	75.4	258	9	US-09-813-453A-49
4	960	74.2	262	9	US-09-813-453A-45
5	862	66.7	233	9	US-09-813-453A-17
6	814.5	63.0	256	9	US-09-813-453A-55
7	771	59.6	255	9	US-09-813-453A-7
8	662	51.2	265	9	US-09-813-453A-4
9	648.5	50.2	250	9	US-09-813-453A-3
10	598	46.2	260	9	US-09-813-453A-51
11	528	40.8	258	9	US-09-813-453A-6
12	527	40.8	219	9	US-09-813-453A-57
13	487.5	37.7	272	9	US-09-813-453A-5
14	483.5	37.4	272	9	US-09-712-363-276
15	437.5	33.8	262	9	US-09-813-453A-8
16	389.5	30.1	246	9	US-09-813-453A-9
17	330.5	25.6	212	9	US-09-813-453A-59
18	322	24.9	257	9	US-09-813-453A-53
19	320.5	24.8	273	9	US-09-813-453A-10

20	275.5	21.3	262	9	US-09-813-453A-11
21	197.5	15.3	244	9	US-09-813-453A-41
22	179.5	13.9	241	9	US-09-813-453A-63
23	138.5	10.7	229	9	US-09-813-453A-12
24	137.5	10.6	257	9	US-09-813-453A-13
25	137	10.6	450	9	US-09-813-453A-39
26	137	10.6	552	9	US-09-813-453A-22
27	136	10.5	552	9	US-09-813-453A-43
28	131.5	10.2	267	9	US-09-813-453A-15
29	121	9.1	248	9	US-09-813-453A-20
30	117.5	8.1	249	9	US-09-813-453A-70
31	114	8.8	249	9	US-09-813-453A-61
32	110	8.5	209	9	US-09-813-453A-21
33	107.5	8.3	242	9	US-09-813-453A-65
34	91	7.0	223	9	US-09-895-913A-74
35	91	7.0	223	9	US-09-813-453A-14
36	91	7.0	223	9	US-09-813-453A-67
37	84	6.5	225	9	US-09-738-626-4158
38	80.5	6.2	958	9	US-09-953-280-43
39	80	6.2	343	9	US-10-060-432-28
40	80	6.2	343	10	US-09-905-173-28
41	79.5	6.1	518	9	US-09-738-626-4769
42	79	6.1	2492	10	US-09-991-258-3
43	79	6.1	4834	9	US-10-097-534-27
44	78.5	6.1	317	9	US-09-533-029-4
45	78.5	6.1	317	9	US-10-286-264-114

ALIGNMENTS

RESULT 1

US-09-813-453A-47
; Sequence 47, Application US/09813453A
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: ANTIBIOTICS
; FILE REFERENCE: OGZ-001
; CURRENT APPLICATION NUMBER: US/09/813,453A
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/227,860
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/667,569
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Bacillus halodurans
US-09-813-453A-47

Query Match 100.0%; Score 1293; DB 9; Length 254;
Best Local Similarity 100.0%; Pred. No. 2.8e-122;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLIVDVGNTVGLGYQDETLVHHWRLATSRQKTEDEYAMTVRSFLFDHAGLQFDIDGI	60
DB	1	MLIVDVGNTVGLGYQDETLVHHWRLATSRQKTEDEYAMTVRSFLFDHAGLQFDIDGI	60
QY	61	VTSVVPVPMFSLQCKKYFHTVPMIIGPKTGLNKNYDNPKEVGADRIYNAVAIEL	120
DB	61	VTSVVPVPMFSLQCKKYFHTVPMIIGPKTGLNKNYDNPKEVGADRIYNAVAIEL	120
QY	121	YGPAIVVDFTATTYCLINEKKQYAGGVIAFGIMISTEALYHRASKLPRIETAKQYV	180
DB	121	YGPAIVVDFTATTYCLINEKKQYAGGVIAFGIMISTEALYHRASKLPRIETAKQYV	180
QY	181	GTNTIDSMOSGIFGYVSQVDGVVYKMKQAASEPKVIATGGTGLAKLIGTSETIDVDSF	240
DB	181	GTNTIDSMOSGIFGYVSQVDGVVYKMKQAASEPKVIATGGTGLAKLIGTSETIDVDSF	240

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

protein - protein search, using sw model

Run on: June 24, 2003, 22:08:46 ; Search time 11.6261 Seconds
(without alignments)
2373.336 Million cell updates/sec

Title: US-09-813-453A-7
Perfect score: 1279
Sequence: 1 MLVLDVGNINVLGIYDGE.....AVERYLEGLRILYRNRE 255

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779.

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB_PEP.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB_PEP.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB_PEP.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB_PEP.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB_PEP.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB_PEP.*
 - 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB_PEP.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB_PEP.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB_PEP.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB_PEP.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB_PEP.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB_PEP.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB_PEP.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1279	100.0	255	9	US-09-813-453A-7
2	834	65.2	258	9	US-09-813-453A-49
3	785.5	61.4	256	9	US-09-813-453A-55
4	783	61.2	262	9	US-09-813-453A-45
5	771	60.3	254	9	US-09-813-453A-47
6	756	59.1	258	9	US-09-813-453A-2
7	650	50.8	233	9	US-09-813-453A-17
8	631.5	49.4	250	9	US-09-813-453A-3
9	629	49.2	265	9	US-09-813-453A-4
10	561	43.9	260	9	US-09-813-453A-51
11	558	43.6	219	9	US-09-813-453A-57
12	554	43.3	258	9	US-09-813-453A-6
13	509.5	39.8	272	9	US-09-813-453A-5
14	505.5	39.5	272	9	US-09-712-363-276
15	414.5	32.4	262	9	US-09-813-453A-8
16	371.5	29.0	273	9	US-09-813-453A-10
17	361.5	28.3	246	9	US-09-813-453A-9
18	338.5	26.5	212	9	US-09-813-453A-59
19	301	23.5	257	9	US-09-813-453A-53

20	266.5	20.8	262	9	US-09-813-453A-11
21	208	16.3	244	9	US-09-813-453A-41
22	172.5	13.5	241	9	US-09-813-453A-63
23	163.5	12.8	249	9	US-09-813-453A-70
24	159	12.4	249	9	US-09-813-453A-61
25	148	11.6	267	9	US-09-813-453A-15
26	143	11.2	460	9	US-09-813-453A-39
27	141	11.0	248	9	US-09-813-453A-20
28	141	11.0	592	9	US-09-813-453A-22
29	139.5	10.9	229	9	US-09-813-453A-12
30	139.5	10.9	592	9	US-09-813-453A-43
31	128	10.0	242	9	US-09-813-453A-65
32	126	9.9	257	9	US-09-813-453A-13
33	100.5	7.9	223	9	US-09-895-913A-74
34	100.5	7.9	223	9	US-09-813-453A-14
35	100.5	7.9	223	9	US-09-813-453A-67
36	94	7.3	610	10	US-09-815-242-11183
37	89	7.0	317	12	US-10-043-238-1
38	89	7.0	317	12	US-10-043-238-3
39	86	6.7	387	10	US-09-925-300-1477
40	83.5	6.5	388	9	US-10-154-251-52
41	83	6.5	4551	9	US-09-793-708-1
42	82	6.4	485	9	US-09-738-626-4551
43	81.5	6.4	714	10	US-09-815-242-10266
44	80.5	6.3	398	10	US-09-815-242-10390
45	80.5	6.3	472	10	US-09-822-863-4

ALIGNMENTS

RESULT 1

US-09-813-453A-7
; Sequence 7, Application US/09813453A
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: ANTIBIOTICS
; FILE REFERENCE: OG2-001
; CURRENT APPLICATION NUMBER: US/09/813,453A
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/227,860
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/667,569
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Geobacter sulfurreducens
US-09-813-453A-7

Query Match	100.0%	Score	1279;	DB	9;	Length	255;
Best Local Similarity	100.0%	Pred. No.	7.3e-120;				
Matches	255;	Conservative	0;	Mismatches	0;	Indels	0;
QY	1	MLVLDVGNINVLGIYDGERLVRDVRVSTDKARTDEYGLINELFRLAGLGDQIRAV	60				
Db	1	MLVLDVGNINVLGIYDGERLVRDVRVSTDKARTDEYGLINELFRLAGLGDQIRAV	60				
QY	61	IISSVVPPLTGVLESLGIFGMRPLVWGFGIKTGMPIQYDNPREVGADEIVNAVAYEK	120				
Db	61	IISSVVPPLTGVLESLGIFGMRPLVWGFGIKTGMPIQYDNPREVGADEIVNAVAYEK	120				
QY	121	YFTSLIIVDFGATTFDYNNRKGECGGGAIAFGVLVISTEALFORASKLPDIDIRPAII	180				
Db	121	YFTSLIIVDFGATTFDYNNRKGECGGGAIAFGVLVISTEALFORASKLPDIDIRPAII	180				
QY	181	ARNTVNSQAGIYGVGLVDEIVYEMKAESKDAPVIATGGLASLAPSKTATVETVY	240				
Db	181	ARNTVNSQAGIYGVGLVDEIVYEMKAESKDAPVIATGGLASLAPSKTATVETVY	240				

THIS PAGE BLANK (USPTO)

senCore version 5.1.6
Copyright (C) 1993 - 2003 CompuGen Ltd.

protein search, using sw.model

June 24, 2003, 22:08:46 ; Search time 11.3982 Seconds
(without alignments)
2373.336 Million cell updates/sec

Title: US-09-813-453A-3

Perfect score: 1257

Sequence: 1 NKRAAFMLLLFLRSVLKVL.....GRRTSLVLATGLAKLIN 250

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	1257	100.0	250	9	US-09-813-453A-3	Sequence 3, Appl1
2	695	55.3	256	9	US-09-813-453A-55	Sequence 55, Appl
3	665.5	52.9	262	9	US-09-813-453A-45	Sequence 45, Appl
4	648.5	51.6	254	9	US-09-813-453A-47	Sequence 47, Appl
5	646.5	51.4	258	9	US-09-813-453A-2	Sequence 2, Appl1
6	631.5	50.2	255	9	US-09-813-453A-49	Sequence 49, Appl
7	629.5	50.1	258	9	US-09-813-453A-17	Sequence 17, Appl
8	623	49.6	233	9	US-09-813-453A-57	Sequence 57, Appl
9	513	40.8	219	9	US-09-813-453A-4	Sequence 4, Appl1
10	498.5	39.7	265	9	US-09-813-453A-4	Sequence 51, Appl
11	460.5	36.6	260	9	US-09-813-453A-51	Sequence 51, Appl
12	428.5	34.1	258	9	US-09-813-453A-6	Sequence 276, App
13	392	31.2	272	9	US-09-712-363-276	Sequence 5, Appl1
14	389	30.9	272	9	US-09-813-453A-5	Sequence 8, Appl1
15	337	26.8	262	9	US-09-813-453A-8	Sequence 9, Appl1
16	333.5	26.5	246	9	US-09-813-453A-9	Sequence 59, Appl
17	327.5	26.1	212	9	US-09-813-453A-59	Sequence 10, Appl
18	282.5	22.5	273	9	US-09-813-453A-10	Sequence 11, Appl
19	262	20.8	262	9	US-09-813-453A-11	

20	229.5	18.3	257	9	US-09-813-453A-53	Sequence 53, Appl
21	169	13.4	241	9	US-09-813-453A-63	Sequence 63, Appl
22	169	13.4	460	9	US-09-813-453A-39	Sequence 39, Appl
23	168	13.4	592	9	US-09-813-453A-22	Sequence 22, Appl
24	168	13.4	592	9	US-09-813-453A-43	Sequence 43, Appl
25	167.5	13.3	244	9	US-09-813-453A-41	Sequence 41, Appl
26	159.5	12.7	248	9	US-09-813-453A-20	Sequence 20, Appl
27	156	12.4	249	9	US-09-813-453A-61	Sequence 61, Appl
28	141	11.2	267	9	US-09-813-453A-15	Sequence 15, Appl
29	136.5	10.9	249	9	US-09-813-453A-12	Sequence 70, Appl
30	123.5	9.8	229	9	US-09-813-453A-13	Sequence 13, Appl
31	121.5	9.7	257	9	US-09-813-453A-13	Sequence 65, Appl
32	108.5	8.6	242	9	US-09-813-453A-65	Sequence 74, Appl
33	102.5	8.2	223	9	US-09-895-913A-74	Sequence 14, Appl
34	102.5	8.2	223	9	US-09-813-453A-14	Sequence 67, Appl
35	102.5	8.2	223	9	US-09-813-453A-67	Sequence 33, Appl
36	86	6.8	872	9	US-10-047-260-38	Sequence 414, App
37	85.5	6.8	542	10	US-09-741-669-414	Sequence 16, Appl
38	84.5	6.7	542	9	US-09-308-207-16	Sequence 1429, A
39	83.5	6.6	856	10	US-09-815-242-11489	Sequence 2, Appl1
40	82	6.5	396	9	US-10-285-074-2	Sequence 4, Appl1
41	82	6.5	396	9	US-10-285-074-4	Sequence 11310, A
42	82	6.5	856	10	US-09-815-242-11310	Sequence 193, App
43	81	6.4	533	9	US-10-108-605-193	Sequence 11, Appl
44	79.5	6.3	607	9	US-10-072-436-11	Sequence 592, Ap
45	79	6.3	854	10	US-09-815-242-5192	

ALIGNMENTS

RESULT 1

US-09-813-453A-3

Sequence 3, Application US/09813453A

Patent No. US20020168681A1

GENERAL INFORMATION:

APPLICANT: Yocum, R. Rogers

APPLICANT: Patterson, Thomas A.

TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF

TITLE OF INVENTION: ANTIBIOTICS

FILE REFERENCE: CG2-001

CURRENT APPLICATION NUMBER: US/09/813,453A

PRIOR FILING DATE: 2001-03-20

PRIOR APPLICATION NUMBER: US 60/227,860

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: 09/667,569

PRIOR FILING DATE: 2000-09-21

NUMBER OF SEQ ID NOS: 77

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 250

TYPE: PRT

ORGANISM: Clostridium acetobutylicum

US-09-813-453A-3

Query Match 100.0%; Score 1257; DB 9; Length 250;
Best Local Similarity 100.0%; Pred. No. 4.9e-120;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	NKRAAFMLLLFLRSVLKVLVDVGNNTVLGIYNDTKLTAEWRLSTDVLSRSDYGIQV	60
DB	1	NKRAAFMLLLFLRSVLKVLVDVGNNTVLGIYNDTKLTAEWRLSTDVLSRSDYGIQV	60
QY	61	MNLFQDKLDPTLVGVIISSVVPNIMTSLHEMKRYKFINPLWVGPIKTNINIKYDP	120
DB	61	MNLFQDKLDPTLVGVIISSVVPNIMTSLHEMKRYKFINPLWVGPIKTNINIKYDP	120
QY	121	KEVGADRIYNAVAHEIYKRSLLIIDFGTATTCFAVRENGDYLGGACGKYSSELEF	180
DB	121	KEVGADRIYNAVAHEIYKRSLLIIDFGTATTCFAVRENGDYLGGACGKYSSELEF	180
QY	181	KAALPRVELIKPAVAICNKTSSQSGIYVYLRQVKYLFELKKNLPLGKATRSLEVL	240
DB	181	KAALPRVELIKPAVAICNKTSSQSGIYVYLRQVKYLFELKKNLPLGKATRSLEVL	240

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: June 24, 2003, 22:08:46 ; Search time 11.307 Seconds
(without alignments)
2373.336 Million cell updates/sec

Title: US-09-813-453A-20
Perfect score: 1248
Sequence: 1 MILEDCGSLIKRWIEGA.....GARIMDLVFGIALACPIE 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues
Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications_AA.*
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
 - 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1248	100.0	248	9	US-09-813-453A-20
2	784.5	63.7	249	9	US-09-813-453A-61
3	789.5	63.3	249	9	US-09-813-453A-70
4	234.5	18.8	592	9	US-09-813-453A-43
5	232.5	18.6	460	9	US-09-813-453A-39
6	231.5	18.5	592	9	US-09-813-453A-22
7	195.5	15.7	244	9	US-09-813-453A-41
8	193	15.5	242	9	US-09-813-453A-65
9	178.5	14.3	267	9	US-09-813-453A-15
10	160	12.8	260	9	US-09-813-453A-51
11	159.5	12.8	223	9	US-09-895-913A-74
12	159.5	12.8	223	9	US-09-813-453A-14
13	159.5	12.8	223	9	US-09-813-453A-67
14	159.5	12.8	250	9	US-09-813-453A-3
15	157	12.6	257	9	US-09-813-453A-53
16	149	11.9	272	9	US-09-813-453A-5
17	147	11.8	272	9	US-09-712-363-276
18	145	11.6	246	9	US-09-813-453A-9
19	144.5	11.6	241	9	US-09-813-453A-63

20	141	11.3	255	9	US-09-813-453A-7	Sequence 7, Appl
21	138.5	11.1	265	9	US-09-813-453A-4	Sequence 4, Appl
22	136	10.9	219	9	US-09-813-453A-57	Sequence 57, Appl
23	136	10.9	258	9	US-09-813-453A-6	Sequence 6, Appl
24	136	10.9	262	9	US-09-813-453A-45	Sequence 45, Appl
25	136	10.9	273	9	US-09-813-453A-10	Sequence 10, Appl
26	133	10.7	258	9	US-09-813-453A-2	Sequence 2, Appl
27	123.5	9.9	258	9	US-09-813-453A-49	Sequence 49, Appl
28	121	9.7	254	9	US-09-813-453A-21	Sequence 21, Appl
29	120.5	9.7	209	9	US-09-813-453A-55	Sequence 55, Appl
30	119.5	9.6	256	9	US-09-813-453A-17	Sequence 17, Appl
31	111	8.9	233	9	US-09-813-453A-59	Sequence 59, Appl
32	107	8.6	212	9	US-09-813-453A-13	Sequence 13, Appl
33	107	8.6	257	9	US-09-813-453A-12	Sequence 12, Appl
34	98.5	7.9	229	9	US-09-813-453A-80-2	Sequence 2, Appl
35	89	7.1	4150	9	US-09-808-880-2	Sequence 4512, Ap
36	88.5	7.1	277	9	US-09-738-626-4512	Sequence 4732, Ap
37	86	6.9	277	9	US-09-738-626-4732	Sequence 253, App
38	85	6.8	298	9	US-09-712-363-253	Sequence 11925, A
39	84	6.7	956	10	US-09-815-242-11925	Sequence 261, App
40	84	6.7	1827	9	US-09-712-363-261	Sequence 11, Appl
41	83.5	6.7	262	9	US-09-813-453A-11	Sequence 8, Appl
42	83	6.7	392	10	US-09-945-825-8	Sequence 4, Appl
43	83	6.7	467	10	US-09-841-880-4	Sequence 12053, A
44	82.5	6.6	251	10	US-09-815-242-12053	Sequence 220, App
45	82.5	6.6	258	9	US-09-712-363-220	

ALIGNMENTS

RESULT 1

US-09-813-453A-20
; Sequence 20, Application US/09813453A
; Patent No. US20020168681A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: ANTIBIOTICS
; FILE REFERENCE: OG2-001
; CURRENT APPLICATION NUMBER: US/09/813,453A
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/227,860
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/667,569
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-813-453A-20

Query Match	100.0%	Score 1248;	DB 9;	Length 248
Best Local Similarity	100.0%	Pred. No. 1.3e-117;	Mismatches 0;	Indels 0;
Matches 248;	Conservative 0;			
QY	1	MILEDCGSLIKRWIEGAARSVAGGLAESDDALVEQLTSQALPVRACRVASRSEQE	60	
Db	1	MILEDCGSLIKRWIEGAARSVAGGLAESDDALVEQLTSQALPVRACRVASRSEQE	60	
QY	61	TSOLVARLEQLFPVSALVASSGKQLAGVRNGYLDYORGLDRWLVAHAHLAKACLVY	120	
Db	61	TSOLVARLEQLFPVSALVASSGKQLAGVRNGYLDYORGLDRWLVAHAHLAKACLVY	120	
QY	121	DLGFAVTSDLVAADGVHLGGYICPGMTLMRSQRLTHTRIRYDDAEARALASLOPQAT	180	
Db	121	DLGFAVTSDLVAADGVHLGGYICPGMTLMRSQRLTHTRIRYDDAEARALASLOPQAT	180	
QY	181	AEAYERGCLLMRGFVREQYAMACLLGPDCEIFLTGDAELVRDELAGATMPDLVFG	240	
Db	181	AEAYERGCLLMRGFVREQYAMACLLGPDCEIFLTGDAELVRDELAGATMPDLVFG	240	

THIS PAGE BLANK (USPTO)

protein search, using sw model

June 24, 2003, 22:24:12 ; Search time 13.838 Seconds
(without alignments)
1947.059 Million cell updates/sec

Title: US-09-813-453A-61
Perfect score: 1269
Sequence: 1 MILEDCGNSFKRWVIRHVA.....QARVVPDLVFLGAMACPLD 249

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	249	9	US-09-813-453A-61
2	849	66.9	249	9	US-09-813-453A-70
3	794.5	62.6	248	9	US-09-813-453A-20
4	245	19.3	592	9	US-09-813-453A-43
5	243	19.1	592	9	US-09-813-453A-22
6	232	18.3	460	9	US-09-813-453A-39
7	210	16.5	265	9	US-09-813-453A-4
8	181	14.3	244	9	US-09-813-453A-41
9	168.5	13.3	242	9	US-09-813-453A-65
10	163	12.8	267	9	US-09-813-453A-15
11	160.5	12.6	272	9	US-09-813-453A-5
12	159	12.5	255	9	US-09-813-453A-7
13	157.5	12.4	272	9	US-09-712-363-276
14	156	12.3	250	9	US-09-813-453A-3
15	154	12.1	258	9	US-09-813-453A-2
16	153.5	12.1	260	9	US-09-813-453A-51
17	144.5	11.4	243	9	US-09-895-913A-74
18	144.5	11.4	223	9	US-09-813-453A-14
19	144.5	11.4	223	9	US-09-813-453A-67

20	142	11.2	262	9	US-09-813-453A-45	Sequence 45, Appl
21	139	11.0	258	9	US-09-813-453A-49	Sequence 49, Appl
22	134	10.6	219	9	US-09-813-453A-57	Sequence 57, Appl
23	131	10.3	256	9	US-09-813-453A-55	Sequence 55, Appl
24	130.5	10.3	257	9	US-09-813-453A-53	Sequence 53, Appl
25	129.5	10.2	233	9	US-09-813-453A-17	Sequence 17, Appl
26	124	9.8	258	9	US-09-813-453A-6	Sequence 6, Appl
27	118	9.3	257	9	US-09-813-453A-13	Sequence 13, Appl
28	118	9.3	273	9	US-09-813-453A-10	Sequence 10, Appl
29	117.5	9.3	246	9	US-09-813-453A-5	Sequence 5, Appl
30	117	9.2	277	9	US-09-738-626-4732	Sequence 4732, Ap
31	114.5	9.0	209	9	US-09-813-453A-21	Sequence 21, Appl
32	114.5	9.0	212	9	US-09-813-453A-59	Sequence 59, Appl
33	114	9.0	254	9	US-09-813-453A-47	Sequence 47, Appl
34	104.5	8.2	241	9	US-09-813-453A-63	Sequence 63, Appl
35	95	7.5	262	9	US-09-813-453A-8	Sequence 8, Appl
36	92.5	7.3	229	9	US-09-813-453A-12	Sequence 12, Appl
37	90	7.1	1832	9	US-10-014-717-4	Sequence 4, Appl
38	89	7.0	218	9	US-09-738-626-6401	Sequence 6401, Ap
39	87.5	6.9	540	9	US-09-712-363-169	Sequence 169, App
40	87.5	6.9	540	9	US-10-267-311-4	Sequence 4, Appl
41	87.5	6.9	540	9	US-10-046-649-4	Sequence 4, Appl
42	87.5	6.9	540	10	US-09-847-6378-6	Sequence 6, Appl
43	87.5	6.9	639	9	US-10-267-311-17	Sequence 17, Appl
44	87.5	6.9	648	9	US-10-267-311-29	Sequence 29, Appl
45	87.5	6.9	690	9	US-10-068-059-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-813-453A-61
; Sequence 61, Application US/09813453A

; Patent No. US20020168681A1

; GENERAL INFORMATION:

; APPLICANT: Yocum, R. Rogers

; APPLICANT: Patterson, Thomas A.

; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF

; FILE REFERENCE: OGZ-001

; CURRENT APPLICATION NUMBER: US/09/813,453A

; CURRENT FILING DATE: 2001-03-20

; PRIOR APPLICATION NUMBER: US 60/227,860

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: 09/667,569

; PRIOR FILING DATE: 2000-09-21

; NUMBER OF SEQ ID NOS: 77

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 61

; TYPE: PRT

; ORGANISM: Pseudomonas putida

US-09-813-453A-61

Query Match 100.0%; Score 1269; DB 9; Length 249;
Best Local Similarity 100.0%; Pred. No. 1,le-116;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MILEDCGNSFKRWVIRHVA	DAVIEGGIVDSOALVAEVAALASVRLTGCRIVSVRSEE	60
DB	1	MILEDCGNSFKRWVIRHVA	DAVIEGGIVDSOALVAEVAALASVRLTGCRIVSVRSEE	60
QY	61	ETDALCALIAQAFVQAKVAHPVREMA	GVNGYDYOGLGMDRWLAALGAFHLAKGACLV	120
DB	61	ETDALCALIAQAFVQAKVAHPVREMA	GVNGYDYOGLGMDRWLAALGAFHLAKGACLV	120
QY	121	IDLGTRAKADFVSADGHEHGGYICPG	MLMRSQSLTRTRRRYRDDASAEALSSLSPO	180
DB	121	IDLGTRAKADFVSADGHEHGGYICPG	MLMRSQSLTRTRRRYRDDASAEALSSLSPO	180
QY	181	TVEAVERGCYLMQGFAYTQLEQARV	LWGEFFVLTGGDAPLVRAALPQARVVDLV	240
DB	181	TVEAVERGCYLMQGFAYTQLEQARV	LWGEFFVLTGGDAPLVRAALPQARVVDLV	240

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

Protein search, using sw model

June 24, 2003, 22:08:46 ; Search time 12.4012 Seconds
(without alignments)
2373.336 Million cell updates/sec

Title: US-09-813-453A-5
Perfect score: 1367
Sequence: 1 MLLAIDVRNTHVTVVGLLSGM.....LRLVFNLEVGORLKTAR 272

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues
Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1367	100.0	272	9	US-09-813-453A-5
2	1363	99.7	272	9	US-09-712-363-276
3	698.5	51.1	265	9	US-09-813-453A-4
4	513	37.5	256	9	US-09-813-453A-55
5	509.5	37.3	258	9	US-09-813-453A-7
6	494.5	36.2	255	9	US-09-813-453A-2
7	487.5	35.7	254	9	US-09-813-453A-47
8	479.5	35.1	258	9	US-09-813-453A-49
9	471.5	34.5	262	9	US-09-813-453A-45
10	431.5	31.6	258	9	US-09-813-453A-6
11	412	30.1	233	9	US-09-813-453A-17
12	404	29.6	262	9	US-09-813-453A-8
13	389	28.5	250	9	US-09-813-453A-3
14	382.5	28.0	260	9	US-09-813-453A-51
15	371	27.1	219	9	US-09-813-453A-57
16	286	20.9	246	9	US-09-813-453A-9
17	274.5	20.1	257	9	US-09-813-453A-53
18	254	18.6	273	9	US-09-813-453A-10
19	209.5	15.3	212	9	US-09-813-453A-59

20	160.5	11.7	249	9	US-09-813-453A-61
21	159	11.6	241	9	US-09-813-453A-63
22	153	11.2	244	9	US-09-813-453A-41
23	152	11.1	592	9	US-09-813-453A-22
24	150.5	11.0	249	9	US-09-813-453A-70
25	149	10.9	248	9	US-09-813-453A-20
26	147.5	10.8	592	9	US-09-813-453A-43
27	146.5	10.7	460	9	US-09-813-453A-39
28	138	10.1	262	9	US-09-813-453A-11
29	115	8.4	338	9	US-09-975-719-47
30	114	8.3	229	9	US-09-813-453A-12
31	112	8.2	242	9	US-09-813-453A-65
32	106	7.8	257	9	US-09-813-453A-13
33	100.5	7.4	264	9	US-09-712-363-183
34	94	6.9	317	12	US-10-043-238-1
35	94	6.8	317	12	US-10-043-238-3
36	92.5	6.8	267	9	US-09-813-453A-15
37	92	6.7	1510	9	US-09-738-626-3707
38	91.5	6.7	209	9	US-09-813-453A-21
39	89.5	6.5	223	9	US-09-895-913A-74
40	89.5	6.5	223	9	US-09-813-453A-14
41	89.5	6.5	223	9	US-09-813-453A-67
42	89	6.5	326	10	US-09-905-176-9
43	89	6.5	337	10	US-09-815-242-12096
44	88	6.4	328	9	US-10-164-433-4
45	87	6.4	789	9	US-09-712-363-244

ALIGNMENTS

RESULT 1
US-09-813-453A-5
Sequence 5, Application US/09813453A
Patent No. US20020168681A1
GENERAL INFORMATION:
APPLICANT: Vocum, R. Rogers
APPLICANT: Patterson, Thomas A.
TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF ANTIBIOTICS
TITLE OF INVENTION: ANTIBIOTICS
FILE REFERENCE: OG2-001
CURRENT APPLICATION NUMBER: US/09/813,453A
CURRENT FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: US 60/227,860
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 09/667,569
PRIOR FILING DATE: 2000-09-21
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 272
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-813-453A-5

Query Match	100.0%	Score	1367	DB	9	Length	272
Best Local Similarity	100.0%	Pred. No.	9.1e-115	Indels	0	Gaps	0
Matches	272	Conservative	0	Mismatches	0		
QY	1	MLLAIDVRNTHVTVVGLLSGMKEHAKYVQOVRIRTESEVTADELALITDGLIGEDSERLTG	60				
Db	1	MLLAIDVRNTHVTVVGLLSGMKEHAKYVQOVRIRTESEVTADELALITDGLIGEDSERLTG	60				
QY	61	TAAISTVPSVLHVRIMLDQYWPSPHVLIEPVRTGIPLLVDNPKVEGADRIYVNCCLAA	120				
Db	61	TAAISTVPSVLHVRIMLDQYWPSPHVLIEPVRTGIPLLVDNPKVEGADRIYVNCCLAA	120				
QY	121	DRFKAAIVYDFGSSICVDVVSARKEFLGGAIPGVQSSDAAARSALRVELARPRS	180				
Db	121	DRFKAAIVYDFGSSICVDVVSARKEFLGGAIPGVQSSDAAARSALRVELARPRS	180				
QY	181	VVGKNTVECMQAGAVFGFAGLVGCLVGRIRREDVSGFVDHVAIVATGHTAPLLPELHT	240				
Db	181	VVGKNTVECMQAGAVFGFAGLVGCLVGRIRREDVSGFVDHVAIVATGHTAPLLPELHT	240				

THIS PAGE BLANK (USPTO)